Experimental Analysis of Successes and Failures in Prosocial and Ethical Behavior

Dissertation

Zur Erlangung des akademischen Grades
doctor rerum politicarum
(Dr.rer.pol.)

vorgelegt dem Rat der Wirtschaftswissenschaftlichen Fakultät
der Friedrich-Schiller-Universität Jena
am 11.02.2015

von Serhiy Kandul (MA)

geboren am 06.11.1984 in: Myroni (Poltawska Obl, Ukraine)

Gutachter: .....
Gutachter: .....
Gutachter: .....

Datum der Verteidigung: .....

Acknowledgements

This work would by no means be possible without continuous support from my family, colleagues, and friends. I am extremely thankful to everyone helping me to work on this thesis.

I would like to specially thank Susanne Büchner for her readiness to resolve dozens of issues of her demanding IMPRS-KITA member.

I am grateful to my first supervisor Prof. Oliver Kirchkamp for helping me with the challenges of the empirical analysis, giving valuable advices, and correcting my mistakes. I thank my second supervisor Prof. Christoph Engel for the inspiring discussions that kept me going further. I deeply appreciate the input of my co-author and friend Matthias Uhl, his remarkable perseverance and dedication.

Open door policy and overall inspiring research atmosphere at the Max Planck Institute perished by Prof. Werner Güth, Prof. Rene Levinsky and Prof. Vittoria Levati deserve my special gratitude.

I thank my IMPRS friends, Magda, Bene, Alex, Mike, Wlad, Henning, Georgiy, Pascal, Carlos, Marco, Marian, Gulnaz, Ayu, Andra, Laura as well as our ESI postdocs Sebastian, Tobie, Mitesh, Sven, Anna and Federica for unforgettable time at the institute.

I deeply acknowledge the support from Sylvia Arnold and Igor Asanov and the input of our HiWis, Johanna, David, Isabel, Maximillian and our lab manager Claudia Zellman.

I thank my wife Lena and my parents Ludmila and Grygoriy for their love and care. I honor Ukrainian people who stand for territorial integrity, peace, democracy, and prosperity of their country.

Contents

Deutsche Zusammenfassung

Der Erfolg von vielen sozialen Interaktionen hängt davon ab, ob Menschen bereit sind, sich ethisch und prosozial zu verhalten. Die Bereitschaft, eigenes Interesse zugunsten anderer zurückzustellen, führt zu Wohlstandsgewinn und Fortschritt der Gesellschaft. Unethisches und antisoziales Verhalten hingegen führen zu Wohlstandsverlusten und Zerstörung.

Diese Doktorarbeit enthält drei Labor-Experimente, in denen die spezifischen Faktoren, die ethisches und pro-soziales Verhalten von Menschen beeinflussen, untersucht werden.

Diese Zusammenfassung beschreibt die Entwicklung des Forschungsrahmens und die wichtigsten Ergebnisse der Doktorarbeit.

Forschungsansätze

Von einkommensverteilungsbasierten Präferenzen bis Selbstbildbewahrung

Jahrelange Wirtschaftsforschung hat die Prämisse vom *Homo öconomicus* in Frage gestellt. In vielen klassischen Experimenten (Diktatorspiel, Vertrauensspiel, Kollektivgutspiel) verhalten sich die Menschen so, dass sie bereit sind, ihre Vergütung zu opfern, um den Zustand der Anderen zu ändern. Dieses Verhalten reizte die Forscher, die sogenannten sozialen Präferenzen in die Wirtschaftstheorie zu intergrieren (Fehr & Schmidt, 1999; Ockenfels & Bolton, 2000; Charness & Rabin, 2002). Laut dieser Modelle kümmern sich die Menschen nicht nur um ihr eigenes Wohlbefinden, sondern auch um ihren Abstand von dem Verdienst der Anderen. *Ungleichheitsaversion* wurde als die Erklärung von solchen 'Anomalien' wie Abgaben

im Diktatorspiel oder Ablehnungen von geringen Beträgen im Ultimatumspiel angeführt. Diese einkommensverteilungsbasierten Modelle wurden später mit der Idee von erwartungsabhängigem Verhalten erweitert: die Menschen nehmen nicht nur die Konsequenzen, sondern auch die Erwartungen der Anderen wahr. Das Paradigma von *Schuldaversion* (Charness & Dufwenberg, 2006; Battigalli & Dufwenberg, 2007) hat es ermöglicht, die Rolle von negativen Emotionen aufgrund der zerstörten Erwartungen der Anderen direkt zu modellieren. Die Forscher haben bemerkt, dass die Möglichkeit, die Erwartungen der Anderen gering zu halten (und trotzdem als 'fair' zu erscheinen), stark das prosoziale Verhalten reduziert (Gueth *et al.*, 1996; Andreoni & Bernheim, 2009; Lazear *et al.*, 2012). Der 'soziale Druck' widersprach der inneren Motivation, sich pro-sozial zu verhalten. Neben 'fairness' (einkommensverteilungsbasierte Modelle), den Erwartungen der Anderen (erwartungsabhängige Modelle) oder sozialem Druck spielen die eigenen Standards der Menschen eine große Role bei sozialem Verhalten, so die Theorien von Selbstbildbewahrung (Akerlof & Kranton, 2010; Mazar *et al.*, 2008; Benabou & Tirole, 2011; Dana *et al.*, 2007). Aus Sicht der Theorie von Selbstbildbewahrung rechtfertigen die Eigenschaften der Umgebung egoistisches Verhalten, so lange das Selbstbild bewahrt ist.

Dana *et al.* (2007) illustriert, wie die Eigenschaften der Umgebung das Resultat der sozialen Interaktion nach der Selbstbildargumentation beinflussen. Im Experiment haben die Teilnehmer die Möglichkeit genutzt, die Konsequenzen ihrer Entscheidungen für die Rezipienten unbekannt zu lassen oder die Verantwortung dafür mit einem anderen Teilnehmer oder mit dem Computer zu teilen. Das Resultat: viel mehr Teilnehmer haben sich für die egoistische Option entschieden.

Abbink & Herrmann (2011) erlaubten den Teilnehmern, die Ausstattung der Anderen zu vernichten. Als die Expertimentatoren den Teilnehmern mitteilten, dass der Computer ungeachtet ihrer Entscheidung die Vergütung der Anderen zerstört, ist die Rate der Vernichtungen stark gestiegen. Der angegebene Gründ dafür ist die Möglichkeit, sich hinter der Situation zu verbergen und 'die moralischen Kosten der Bosheit' zu reduzieren.

Die Wirkung von situativen Rechtfertigungen hat den Forschern geholfen, die gesammelten Nachweise von Großzügigkeit der Menschen mit antisozialem Verhalten (wie Neid-gesteuerte Vernichtung (Casal *et al.*, 2012), antisoziale Bestrafung in Kollektivgutspielen (Thöni, 2014) oder Bosheit und Wohlbefinden reduzierende Sabotage (Harbring & Irlenbusch, 2009)) zu versöhnen.

Selbstbild und ethisches Verhalten

Ethisches Verhalten, das meistens als Reaktion auf die Möglichkeit von der Lüge zu profitieren modelliert wurde, hat die experimentelle Wirtschaft mit zwei wichtigen Musterbeispielen untersucht: Sender-Rezipient Spiel (Gneezy, 2005), wobei man die individuelle Lüge beobachten kann, und 'die-under-cup' Experimente (Fischbacher & Föllmi-Heusi, 2013) mit der Möglichkeit, die aggregierte Verbreitung der Lüge zu messen. Das übliche Ergebnis in dieser Literatur ist das Phänomen der 'Lügenaversion', die Widerwillgkeit, die Gewinne mithilfe der Lüge zu maximieren. Als Motivation für diese Widerwilligkeit wurde Rücksicht auf die Anderen genannt: die Menschen lügen nicht, um den Anderen nicht zu schaden. Gneezy (2005) variiert den Verlust durch die Lüge und beobachtet, dass die Rate der Lüge sinkt, je höher die Verluste der Betroffenen sind. Die anderen Studien ergaben allerdings, dass die Widerwilligkeit zum Lügen

sogar in "win-win" Situationen vorkommt (Erat & Gneezy, 2012). Diese Ergebnisse haben die Motivation, die Wahrheit zu sagen, um den Anderen nicht zu schaden, in Frage gestellt. Erwartungsbasierte Erklärung im Sinne von 'Schuldaversion' wurde auch vorgeschlagen, um den Umgang der Menschen mit der Lüge zu verstehen. Das heißt, die Menschen lügen nicht, weil die Anderen erwarten, dass sie die Wahrheit sagen. Obwohl die 'Second-order beliefs' sich als relevant für die Entscheidung zu lügen erwiesen haben (Battigalli *et al.*, 2013), genügt 'Schuldaversion' nicht, die Widerwilligkeit zu Lügen in den anderen Experimenten zu erklären (Vanberg, 2008). Deswegen vermuteten die Forscher, dass die Tat der Lüge selbst mit psychologischen Kosten verbunden ist (Kartik, 2009).

Im Gegenteil zur Wirtschaft, basiert die Erforschung von unethischem Verhalten in der Psychologie auf der Selbstbildsbewahrung. Shalvi *et al.* (2011) demonstriert, wie die situationsbedingten Entschuldigungen die Lüge verbreiten können. In dem 'die-under-cup' Experiment sollten die Teilnehmer das Ergebnis des Würfelns berichten, um ihren Gewinn zu bestimmen. Die gemessene Rate der Lüge war viel höher in der Bedingung, in der die Teilnehmer mehrfach würfeln durften. Obwohl die Teilnehmer informiert wurden, dass nur der erste Versuch für die Auszahlung gilt, neigten sie dazu, ihre Lüge damit zu rechtfertigen, dass sie eine größere Zahl bei den irrelevanten Versuchen gewürfelt haben.

Mazar *et al.* (2008) analysiert den Effekt von moralischen Erinnerungen auf die Rate der Lüge. In ihrem Experiment konnten die Teilnehmer selbst ihre Leistung bei der Aufgabe berichten ohne bzw. mit einer Erinnerung, sich moralisch zu verhalten, bekommen zu haben. Die Autoren der Studie schlussfolgerten, dass die Erinnerungen an die moralischen Standards, die Ehrlichkeit erhöhen.

Diese Ergebnisse aus prosozialem und ethischem Bereich unter-

stützen die Annahme, dass die Menschen ihre Standards und die Bewertungen von Selbstbild an den Kontext der Situation anpassen können (Festinger, 1962; Konow, 2000; Matthey & Regner, 2011).

Die Identifizierung von den spezifischen Eigenschaften der Umgebung, den situationsbedingten Entschuldigungen, die das Niveau von ethischem und prosozialem Verhalten beinflussen, ist das große Forschungsziel dieser Doktorarbeit.

Prosoziales und ethisches Verhalten im dynamischen Kontext

Ein Großteil der experimentellen Forschung von prosozialem und ethischem Verhalten fokusiert sich auf die Motivation der Menschen in einer vorgegebenen feststehenden Umgebung. Die wachsende Literatur in moralischer Psychologie zeigt allerdings die Wichtigkeit, das menschliche Verhalten im Laufe der Zeit zu untersuchen. Früher wurden die Menschen als konsistent betrachtet, was ihre moralischen Standards oder inneren Normen betrifft. Die Konsistenz liegt in Verfahren für soziale Zustimmung, wie z. B. 'foot-in-the-door' Paradigma (Freedman & Fraser, 1966), wobei kleine prosoziale Aktivität die größere prosoziale Aktivität erzeugt. In vielen Experimenten erzeugte die vergangene (un)ethische Aktivität die Konsistenz in der nächsten Aufgabe (Gino *et al.*, 2010; Cojoc & Stoian, 2014; Gneezy *et al.*, 2012a). Mehrere neue Experimente haben eine alternative Neigung vorgestellt: ausgleichendes Verhalten oder 'balancing' (Nisan & Horenczyk, 1990; Meritt *et al.*, 2010; Mazar & Zhong, 2010). Laut dieses Ansatzes kann eine schlechte (gute) Tat in der Vergangenheit mit einer guten (schlechten) Tat in der Zukunft kompensiert werden. Diese Tendenz hat sich in den verschiedenen Bereichen wie Konsum, politischer Diskriminierung,

oder Spenden erwiesen (Kahn & Dhar, 2006; Monin & Miller, 2001; Sachdeva *et al.*, 2009; Gneezy *et al.*, 2012b).

Meritt *et al.* (2010) erklärt zwei mögliche Mechanismen von ausgleichendem Verhalten durch die Änderungen im Selbstbild der Menschen. Laut der Idee von 'moralischem Kredit', verdienen vergangene moralische Aktivitäten einen Kredit, den man später verwenden kann, wenn man sich in einer moralisch umstrittenen Aktivität engagiert. Vergangene unmoralische Aktivitäten schaden wiederum dem Selbstbild und verdienen moralische Schulden, die man mit einer späteren moralisch überlegeneren Aktivität kompensiert. In beiden Fällen erkennt die Person die entstehenden Inkonsistenz. Der Mechanismus der 'moral credentials' hingegen wirkt so, dass die für das Selbstbild verdienten Prämien den Menschen helfen, die Unklarheit über die Sittlichkeit der nächsten Aktivität aufzulösen, ohne das Gefühl, etwas Unakzeptables getan zu haben.

In dieser Doktorarbeit wurde Nutzen gezogen aus der Verbindung zwischen den vergangenen und zukünftigen Aktivitäten und die Experimenten wurden so gebaut, dass der dynamische Charakter des entstehenden prosozialen und ethischen Verhaltens weiter untersucht werden kann.

Rahmen und die hauptsächlichen Ergebnisse der Studien

Der Einfluss von ex-post Information auf die Spenden in einem Diktatorspiel mit einer unvollständigen Information

Kapitel 2 dieser Doktorarbeit (basierend auf einem 'single-author' Projekt) bereichert die Diskussion über das Selbstbild im proso-

zialen Verhalten, indem der Effekt von ex-post Informationen auf Spenden in einem Diktatorspiel untersucht wurde. Das Experiment in diesem Kapitel wurde so aufgebaut, um die folgende Situation zu modellieren. Geber haben oft keine Möglichkeit, ex-ante zu identifizieren, ob die Rezipienten, die um Hilfe bitten, in der Realität in Not sind. In dieser Situation können die Geber ihre Spenden reduzieren, weil sie befürchten, dass ihre Hilfe missbraucht wird. Das Resultat: die Rezipienten, die wirklich in Not sind, bekommen weniger. Im Experiment manipuliert man die ex-post Information über die Ausstattung der Rezipienten, die die Geber erhalten, um mehr über das Verhalten des Gebers zu entdecken. Einerseits kann die Möglichkeit, unwissend zu bleiben, die Spenden erhöhen, wobei die Geber die psychologischen Kosten, dem 'Falschen' zu helfen, vermeiden können (positive Seite von Unwissenheit). Andererseits kann keine Information den Gebern helfen, ihre Spenden zu reduzieren und ein positives Selbstbild zu behalten (negative Seite von Unwissenheit). Daraufhin erwarten wir, dass ex-post Infromation die Spenden beeinflusst und dass die Geber die Möglichkeit nutzen, nicht zu wissen, was in der Realität die Ausstattung der Rezipienten ist.

Aufbau des Experiments Ich benutze das modifizierte Diktatorspiel mit Anfragen und unvollstandiger Information über die Ausstattung des Rezipienten. Am Anfang des Experiments machen die Teilnehmer eine 'real-effort' Aufgabe. Für diese Aufgabe bekommt die Hälfte der Teilnehmer ('Diktatoren') eine sichere Ausstattung von 10 EUR. Der Verdienst der anderen Hälfte der Teilnehmer ('Rezipienten') ist mit gleicher Wahrscheinlichkeit entweder niedrig (0 EUR) oder groß (6 EUR). Diktatoren wissen ex-ante nicht wie viel der Rezipient hat. Der Rezipient bittet um Hilfe, indem er eine

Anfrage an den Diktator schickt[1]. Die Diktatoren bestimmen den Betrag, den sie ausgeben wollen, für jede mögliche Anfrage des Rezipienten. Der Computer fügt die entsprechenden Entscheidungen von Diktator und Rezipient zusammen. Als Manipulation wechsle ich die Information über die Ausstattung des Rezepienten, die die Diktatoren nach ihrer Entscheidung bekommen. In Treatment 'Info' erfahren die Diktatoren die Ausstattung der Rezipienten; in Treatment 'No_info' erfahren die Diktatoren die Ausstattung der Rezipienten nicht; in Treatment 'not-to-know' konnten sich die Diktatoren (bevor sie eine Anfrage bekommen und entscheiden wie viel Geld sie ausgeben) entscheiden, ob sie die Ausstattung des Rezipienten wissen wollen. Die ersten zwei Bedingungen mit exogener Information wurden 'within-subjects' durchgeführt. Die Bedingung mit endogener Unwissenheit wurde in getrennten Sessions durchgeführt. Die Teilnehmer agierten in zwei Perioden mit perfektem 'stranger-matching' und keinen Rückmeldungen zwischen den Perioden. Ich habe insgesamt 6 Sessions im Oktober 2014 im Labor des Max-Planck-Instituts für Ökonomik durchgeführt. 188 Teilnehmer haben an dem Experiment teilgenommen.

Hauptsächliche Ergebnisse:

- Resultat 1: Diktatoren im 'not-to-know' Treatment haben signifikant weniger ausgegeben; kein Unterschied zwischen den Bedingungen mit exogener Manipulation der Information;
- Resultat 2: Manche Diktatoren entscheiden sich, die Ausstattung des Rezipienten nicht zu wissen. 25%-46% 'not-to-know' Entscheidungen zeigen, dass eine signifikante Minderheit von Dikatoren sich entscheidet, unwissend zu bleiben.

[1]Der Rezipient kann nach jedem Betrag zwischen 0 und 5 EUR fragen.

- Resultat 3: Die Diktatoren, die unwissend bleiben, geben weniger aus als diejenigen, die die Ausstattung des Rezipienten erfahren wollen.

Ich schlussfolgere, dass die ex-post Information eine wichtige Rolle für prosoziales Verhalten spielt. Gleichermaßen zur psychologisch nachteiligen ex-ante Information, die Möglichkeit, ex-post unwissend zu bleiben, reizt Egoismus an und unterstützt größere Ungleichheit.

Gegenwärtiger und zukünftiger Effekt von Delegation der Lüge

Kapitel 3 der Doktorarbeit (basierend auf einem gemeinsamen Projekt mit Prof. Oliver Kirchkamp) bereichert die Studien von der Verbindung zwischen prosozialen und ethischen Verhalten, indem die gegenwärtigen und zukünftigen Effekte von Delegation der Lüge untersucht wurden. Unsere Intuition in diesem Projekt stammt aus dem Unterschied zwischen direkten und indirekten Übeltaten in moralischer Psychologie. Wir haben angenommen, dass die Möglichkeit zu delegieren, das Phänomen der 'Lügenaversion' stark reduzieren kann. Wir testen, wie die Möglichkeit, die Lüge zu delegieren, die zukünftige ausgleichende Aktivität beeinflusst (die Tendenz, die von der 'moralischen Bilanz' Literatur vorgebracht wurde). Wir haben erwarten, dass die Delegation der Lüge den Bedarf, die Übeltat zu kompensieren, vermindern kann.

Aufbau des Experiments In dem Experiment spielen die Teilnehmer zu dritt ein Sender-Rezipient Spiel und ein Diktatorspiel. In dem Sender-Rezipient Spiel agieren die Teilnehmer in Dreiergruppen: zwei Sender ('Spieler 1' und 'Spieler 2') und ein Rezipient

('Spieler 3'). Der Computer verteilt die Preisgelder für die Sender (8 EUR) und für den Rezipient (4 EUR) zwischen fünf virtuellen Kästen und ordnet jede Gruppe einer der zwei Bedingungen zu. Wir vergleichen zwei Bedingungen (within-subjects): 'Kein Konflikt' und 'Konflikt'. Unter der 'Kein Konflikt' Bedingung platziert der Computer die Preisgelder für die Sender und für den Rezipient in den gleichen Kasten, unter der 'Konflikt' Bedingung sind die Preisgelder für die Sender in einem Kasten und das Preisgeld für den Rezipient in einem anderen Kasten platziert. Nachdem die Sender die Verteilung der Preisgelder erfahren haben, werden sie gefragt, welchen Kasten sie dem Rezipienten empfehlen: 'Dein Preisgeld ist im Kasten x' und zu entscheiden, ob sie *selber* die Nachricht schicken oder lieber *delegieren* wollen. Nachdem beide Sender ihre Entscheidungen getroffen haben, wählt der Computer einen Sender aus und führt die entsprechenden Entscheidungen durch. Falls der ausgewählte Sender 'delegiert', wird die vom anderen Sender empfohlene Kastennummer an den Rezipienten übermittelt. Falls der ausgewählte Teilnehmer 'selber' ausgewählt hat, wird die von diesem Sender empfohlene Kastennummer an den Rezipienten übermittelt. Die Rezipienten erhalten dann die entsprechende Nachricht (sie werden informiert, ob die Delegation stattfindet) und wählen einen Kasten, den sie öffnen wollen. Die Entscheidung des Rezipienten bestimmt die Auszahlung der Teilnehmer in dem Spiel.

Nach dem Sender-Rezipient Spiel werden die Teilnehmer neuen Dreiergruppen zugeteilt. In dem Diktatorspiel werden die Sender gefragt, wie viel von ihrem Verdienst im Sender-Rezipient Spiel sie an den Rezipienten geben würden.

Wir erwarteten, dass die Rate der Delegtion in der 'Konflikt' Bedingung höher ist; dass die Sender, die selbst die Wahrheit sagen, dazu neigen, mehr zu delegieren als die Lügner; dass die lügenden

Sender mehr Geld im Diktatorspiel ausgeben, um ihre Übeltat zu kompensieren, und dass die Teilnehmer, die delegieren, aufgrund der niedriger Verantwortung weniger Geld ausgeben als diejenigen, die direkt lügen.

Wir haben 7 Sessions im Labor des Max-Planck-Instituts für Ökonomik in Jena in November-Dezember 2013 durchgeführt. Insgesamt 204 Teilnehmer haben an dem Experiment teilgenommen.

Hauptsächliche Ergebnisse:

- Resultat 1: Die Häufigkeit der Delegation höher in der "Konflikt"Bedingung ist.
- Resultat 2: Unter den delegierenden Sendern entscheiden sich mehr, die Wahrheit zu sagen.
- Resultat 3: Das Spenden in der 'Konflikt' Bedingung ist höher.
- Resultat 4: Die delegierenden Sender geben mehr Geld als die direkt lügenden Sender aus.

Wir schlussfolgern, dass die Möglichkeit zu delegieren die 'Lügenaversion' vermindern kann. Falls die Sender, die sonst die Wahrheit sagen, delegieren, ist die Rate der entstehenden Lüge höher. Im Widerspruch zu unseren Erwartungen geben die delegierenden Sender mehr Geld aus als die Lügner. Wir schlagen vor, dass die Möglichkeit zu delegieren als Selektion-Device gilt: die Sender mit höheren psychologischen Kosten der Lüge delegieren und benutzen die Gelegenheit, ihre Übeltat zu kompensieren; direkt lügende Sender mit niedrigeren psychologischen Kosten der Lüge haben keinen Bedarf, ihre vergangenen Übeltaten auszugleichen. Die direkte Analyse der Verbindung zwischen der Verantwortung für die Übeltat und dem

Willen diese auszugleichen, ist ein viel versprechendes Thema für zukünftige Studien.

Effekt von Aktivität der 'dritten Person' auf moralisches Verhalten

Kapitel 4 dieser Doktorarbeit (basierend auf einem gemeinsamen Projekt mit Dr. Matthias Uhl) bereichert die Diskussion über ethisches und prosoziales Verhalten in einem dynamischen Kontext, indem der Effekt von der Aktivität der 'dritten Person' auf die Tendenz, die vergangenen Taten auszugleichen, untersucht wurde. Es gibt zwei alternative Hypothesen hinsichtlich der Richtung vom Zusammenhang zwischen vergangenen und zukünftigen moralischen Aktivitäten: laut 'moralischer Bilanz' sind vergangene und zukünftige ethische Taten negativ korreliert, laut 'moralischer Konsistenz', sind vergangene und zukünftige ethische Taten positiv korreliert. Die Forschung hat mehrere Faktoren verdeutlicht, die konsistent oder kompensierendes Verhalten erzeugen: die monetären Kosten der vergangenen Taten, das Niveau von konzeptualer Abstraktion (jüngere vs. alte Taten), und der Mind-set (Konsequenzen- oder regelbasierte Neigung). In diesem Projekt untersuchen wir, welche Rolle der Mind-set der Menschen für den Link zwischen eigenen Aktivitäten und den Taten der dritten Person spielt. Wir erwarten, dass die konsequenzenorientierten Menschen die Aktivität der Anderen als Substitut wahrnehmen und ihre Großzügigkeit entsprechend anpassen. Die regelorientierten Menschen, im Gegenteil, nehmen die Tat der Anderen als Ergänzung. Wir testen diese Annahme in einem Laborexperiment.

Aufbau des Experiments Das Experiment besteht aus drei Teilen: Recall-task, Diktatorspiel, und die 'real-effort' Aufgabe. In der Erinnerungsphase werden die Teilnehmer aufgefordert, sich an eigenes jüngeres ethisches oder unethisches Verhalten zu erinnern (between-subjects). Zwischen der Erinnerung und dem Diktatorspiel erfahren die Teilnehmer, wie viel Mühe sie sich in der letzten Aufgabe geben müssen. Die Mühe können niedrig (nur wenige Tabellen korrekt lösen) oder groß (viele Tabellen korrekt lösen) sein. In der Baseline legt der Computer die Mühe fest. Im 'dritte Person' Treatment sind die Mühen von einem anderen Teilnehmer zu bestimmen. Am Ende des Expriments werden alle Teilnehmer gefragt, das Trolley-Dilemma zu beantworten. Diejenigen, die es akzeptabel finden, fünf Menschen zu retten und eine Person zu opfern, werden als konsequenzenorientierte Personen behandelt. Die Teilnehmer, die es unakzeptabel finden, werden als regelorientierte Menschen klassifiziert. Wir erwarteten, dass die Erinnerung an eine unethische Aktivität die konsequenzenorientierten Teilnehmer anreizt, mehr Geld auszugeben (ausgleichendes Verhalten) und regelorientierte Teilnehmer dazu neigen, weniger auszugeben (konsistentes Verhalten). Wir untersuchen, wie die negative (viele Tabellen) oder positive (wenige Tabellen) Tat von einer dritten Person diese Tendenzen beeinflusst.

Wir haben 7 Sessions im August 2014 im Labor des Max-Planck-Instituts für Ökonomik in Jena durgeführt. Insgesamt haben 218 Teilnehmer an dem Experiment teilgenommen.

Hauptsächliche Ergebnisse:

- Resultat 1: Unsere Manipulation erzeugt konsistentes Verhalten. Die Teilnehmer in der 'unethischen' Gruppe sind weniger

großzügig als die Teilnehmer in der 'ethischen' Gruppe. Der Effekt ist stärker für die konsequenzenorientierten Teilnehmer.

- Resultat 2: Konsequenzenorientierte Teilnehmer reagieren eigentlich nicht auf die Tat der dritten Person, aber die regelorientierten Teilnehmer verhalten sich entsprechend. 'Unethische' regelorientierte Teilnehmer verteilen signifikant weniger, wenn sie die netagive Tat der dritten Person beobachten.

Wir schlussfolgern, dass der Unterschied zwischen konsequenzen- und regelorientierte Menschen wichtig ist, um das ethische Verhalten der Menschen zu verstehen. Wir spekulieren, dass die Tat der anderen als Beispiel für die regelorientierten Menschen wirkt.

List of Figures

List of Tables

Chapter 1

Introduction

> The only difference between the
> saint and the sinner
> is that every saint has a past,
> and every sinner has a future
>
> Oscar Wilde

The successfulness of many social interactions depends on people's readiness to behave ethically and pro-socially. The willingness to sacrifice one's own interest in the interest of others leads to welfare gains and social progress. Unethical and antisocial behavior, in contrast, results in unrecoverable welfare losses and destruction.

This thesis includes three laboratory experiments investigating specific factors that affect the degree of people's ethicality and prosociality.

This chapter describes the development of the research framework and the summary of the main findings in the thesis.

1.1 The research framework for experimental studies of pro-social and ethical behavior

From outcome-based preferences to self-image concerns

Decades of experimental research in economics brought a large body of evidence against the pure money maximizing hypothesis of *homo economicus*. In a series of stylized experiments (dictator, trust, public good games) people often sacrifice their resources to change the well-being of others. In late 90s and early 2000s,Fehr & Schmidt (1999); Ockenfels & Bolton (2000); Charness & Rabin (2002) introduced other-regarding concern and social preferences into standard utility model. In these new models, people derive utility not only from their own monetary income but also from their absolute or relative distance from others. Emerging *inequality aversion* was proposed to explain a broad range of behavioral "anomalies" such as giving positive amounts in dictator games or rejecting low offers in ultimatum-like settings. Pure distributional preferences in terms of monetary outcomes were later augmented with the idea of belief-dependent behavior: people behave pro-socially since they want to meet expectations and not only to change the well-being of others.

Guilt aversion paradigm (Charness & Dufwenberg, 2006; Battigalli & Dufwenberg, 2007) as a specific type of belief-dependent behavior in psychological games, allowed researchers to directly model the role of negative emotions from hurting the expectations of others. Originally, the desire to meet expectation of others was suggested to explain rewarding trust by keeping promises in a principal-agent setting. Later *guilt aversion* was shown successful in understanding behavior in a larger class of dictator and ultimatum-like games. In particular, lower observability, possibility to appear fair

or to reduce the expectations of others dramatically reduces pro-social behavior (Gueth *et al.*, 1996; Andreoni & Bernheim, 2009; Lazear *et al.*, 2012). The "social pressure" formed by others' expectations behind generosity challenged the assumption of intrinsic motivation to behave pro-socially. The pressure to behave pro-socially, however, does not necessarily originate from external, social forces, as the literature on self-image suggests.

According to self-image theories, in addition to fairness in outcome-based models, expectations of others and emotions in belief-dependent models or social-image concerns, people's own behavioral standards, their identity or self-image determine the degree of their pro-sociality (Akerlof & Kranton, 2010; Mazar *et al.*, 2008; Benabou & Tirole, 2011; Dana *et al.*, 2007). In principle, people would like to keep a positive self-image by adhering to their internal behavioral standards. However, specific features of the environment may allow them to justify deviations from their own internal norms and still maintain their self-image. In other words, 'situational excuses' can release people from their internal norms and make them pursue a self-interest action they would otherwise consider inappropriate.

Dana *et al.* (2007) illustrates the kind of environmental features affecting the outcomes of social interaction based on the self-image reasoning. In their experiment, participants in the role of dictators exploited the opportunity to remain ignorant about the consequences of their actions for the other player or to share responsibility for their decision with another dictator or with a computer.

Abbink & Herrmann (2011) allowed participants to destroy part of the endowment of others. In one of the treatments, the experimenter informed the participants that the computer will with certain probability reduce the pay-off of their counterpart irrespective of what they decide. This possibility to hide behind the nature

or to 'reduce the moral costs of nastiness' doubled the rate of destruction.

The effectiveness of such self-image tolerant excuses helped researchers to at least partially reconcile accumulated evidence about people's generosity with otherwise puzzling findings of antisocial behavior like envy-related destruction (Casal *et al.*, 2012), antisocial punishment in public good games (Thöni, 2014), spite and welfare-reducing sabotage (Harbring & Irlenbusch, 2009).

Self-image and ethical behavior

Ethical behavior, usually modeled as a reaction towards the possibility to gain by lying, was investigated in the experimental economics literature through two major paradigms: sender-receiver games (Gneezy, 2005) and die-under-cup experiments (Fischbacher & Föllmi-Heusi, 2013). In a sender-receiver game, participants with more information about the state of the world can increase their payment by sending untruthful message to their counterparts. In this case, experimenter can observe individual lies. In die-under-cup experiments, participants are typically asked to self-report the outcome of a die roll and are rewarded for a higher number. In this case, by comparing the expected and reported averages the experimenter can assess the prevalence of lying on the aggregate level.

The common finding in these experiments is the phenomenon of *lying aversion* or the reluctance to maximize the benefits through lying. Other-regarding preferences were proposed as a natural driving force behind the unwillingness to lie: people do not always lie since they do not want to harm others. Gneezy (2005) varies the harm to others due to lie and observes that lying rate shrinks with

the losses to the affected party. Other studies, however, provide evidence of the reluctance to lie even in 'win-win' interactions (Erat & Gneezy, 2012) challenging the importance of other-regarding concern in decisions to lie. Belief-dependent justifications in the spirit of guilt-aversion models have also been proposed to explain lying aversion: people do not want to lie since others expect them to tell the truth. Although second-order beliefs are shown to be correlated with the incidence of lying (Battigalli *et al.*, 2013), guilt-aversion fails to explain the reluctance to lie observed in other experiments (Vanberg, 2008). Similarly to the domain of pro-social behavior, outcome-based preferences or expectations of others do not suffice to fully capture the scope of unethical behavior. These observations lead researchers to claim that lying per se induces costs which truth-tellers are trying to avoid (Kartik, 2009).

In contrast to behavioral economics, research on ethical behavior in psychology appeals to self-image maintenance. In the search of situational excuses Shalvi *et al.* (2011) demonstrate how the availability of favorable counterfactuals can enhance lying. In their die-under-cup experiment, participants had to report the outcome of the die roll which defined their payment. The estimated rate of lying was substantially higher in multiple-roll treatment where participants had the chance to roll the die several times. Although participants were informed that only the first roll counts for payment, they tend to justify their lie by the fact they rolled a higher number in pay-off irrelevant trials (justification by counterfactuals). Mazar *et al.* (2008) analyzes the effects of ethical reminders on the level of cheating. In their experiment, participants could self-report their performance on a task with or without being reminded of the ethical norm. The authors conclude that reminders (by appealing to one's moral standards) substantially increase honesty.

These insights in pro-social and ethical domains suggest that people can adapt their standards of behavior or beliefs about themselves to the context of the situation (Festinger, 1962; Konow, 2000; Matthey & Regner, 2011).

The identification of the specific features of the environment, situational excuses, that affect the degree of people's pro-sociality and ethicality is the broad research goal of this thesis.

Pro-social and ethical behavior in a dynamic context

Most of the experimental research of pro-social and ethical behavior focuses on motivations of people in a given static environment. However, a growing body of literature in moral psychology emphasizes the importance of looking at the developments of people's actions over time. Traditionally, people are believed to behave consistently as regards their moral standards or internal norms. In the well-known social compliance technique called "foot-in-the door" (Freedman & Fraser, 1966), agreement to requests for small pro-social actions (e.g. signing a petition) increases the likelihood of their more substantial pro-sociality (e.g. real donations) in the future (Schwarzwald *et al.*, 1983). In more recent experiments it has been shown that previous (un)ethical action induces consistency in the next task (Gino *et al.*, 2010; Cojoc & Stoian, 2014; Gneezy *et al.*, 2012a).

A cohort of recent experimental studies have put forward an alternative tendency: striving for 'compensatory' or 'moral balancing' behavior (Nisan & Horenczyk, 1990; Meritt *et al.*, 2010; Mazar & Zhong, 2010). According to this view, doing good (bad) in the past might be compensated by a bad (good) action in the future. In Mazar & Zhong (2010) experiment participants who purchased in

a green store cheated more in a subsequent task compared to participants from a conventional online store. In their interpretation, behaving ethically in the first stage 'licenses' deviation downwards in a different domain. Exploring 'moral licensing' effect, Kahn & Dhar (2006) showed that participants who volunteered for a public project were more likely to buy luxury vs. conventional product afterwards; Sachdeva *et al.* (2009) observed more selfishness from participants primed with positive stories about themselves; Monin & Miller (2001) noticed that expressing justice increases the likelihood of future discriminatory decisions. This recent evidence of socially undesirable inconsistency in moral behavior complemented 'moral cleansing' (transgression-compliance effect), a tendency to compensate wrongdoing with subsequent better action, observed in other studies (Jordan *et al.*, 2011; Gneezy *et al.*, 2012b).

Meritt *et al.* (2010) discuss two possible mechanisms behind compensatory behavior, moral licensing and moral cleansing, by appealing to changes in the self-image. According to moral credit theory, previous moral action earns moral credits that a person then spends by engaging in a morally questionable behavior. Previous immoral action, in turn, damages self-image and earns moral debts that a person then seeks to repair by subsequently engaging into a morally superior activity. In both cases, a person is aware of the resulting inconsistency. According to moral credentials theory, earning bonuses to self-image helps people to resolve uncertainty about the ethicality of the subsequent action without realizing of doing anything morally inferior.

The reconciliation of the evidence for consistent and compensatory behavior and identification of necessary conditions under which one or the other tendency dominates is the current research venue in moral dynamics.

Gneezy *et al.* (2012a) manipulates the costs of the pro-social action and argues that the lower the costs behind pro-sociality, the more is the likelihood of moral licensing, whereas higher costs (by providing a reliable signal of one's identity) induce consistency. Clot *et al.* (2013) further corroborate this intuition. In their experiment, financial rewards for volunteering reduced the scope of licensing.

Conway & Peetz (2012) appeal to level construction theory proposed by Trope & Liberman (2003) and suggest that the level of conceptual abstraction (distant vs. recent actions) defines whether consistent or compensatory behavior emerges. Parallel to goal pursuit research by Fischbach *et al.* (2006), distant (im)moral actions enhance commitment to a goal of maintaining positive self-image and result in consistency, whereas recent (im)moral actions signal progress towards a goal and allow to reduce moral strivings in the future[1].

Cornellissen *et al.* (2013) propose that the mind-set of people (outcome vs. rule-based) explains the direction of the link between past and future moral actions. In their experiment, participants classified as outcome-based (willing to pull the leverage in a trolley dilemma) exercised moral balancing by giving more in a dictator game after recalling an unethical action), whereas rule-based participants (unwilling to pull the leverage in a trolley dilemma) exhibited consistency by giving less in a dictator game after recalling an unethical action.

The experiments included in this thesis focus on the link between one's past and future actions and are designed to further explore the dynamic nature of the resulting pro-social and ethical behavior.

[1]See also Jordan *et al.* (2011) for the discussion of how moral behavior relates to goal pursuit.

1.2 The scope and the summary of the main findings of the thesis

1.2.1 The effect of ex-post information disclosure on giving in a dictator game with incomplete information

Chapter 2 of this thesis (based on a single-author project) contributes to the discussion of self-image concerns in pro-social behavior by investigating the effect of ex-post information disclosure on a giving in a dictator game. The experiment described in this chapter is modeled to reflect the following situation. Donors approached with the donation request often have no possibility to ex-ante identify if the recipient is in a true need or not. In such an environment, donors might reduce their willingness to donate to avoid the possibility that their help is misdirected. As a result, the recipients with a true need do not receive necessary help. I manipulate the ex-post information disclosure of recipient's endowment to get more insights about donor's behavior. The opportunity to remain ignorant can help donors to increase their transfers by avoiding the psychological costs of helping "the wrong' recipient (positive effect of ignorance). However, no information disclosure can help donors to reduce their transfers while maintaining positive self-image (negative effect of ignorance). I hypothesize that ex-post information disclosure affects dictator's giving and that participants employ the possibility to remain ignorant about recipient's endowment.

Experimental Design I employ a modified dictator game with requests and incomplete information about recipient's endowment.

At the beginning of the experiment participants perform a real-effort task. For completion of this task half of the participants

receive a sure endowment of 10 EUR (Dictators), the earnings of the other half of the participants (Recipients) are either low (0 EUR) or high (6 EUR) with equal probability. Dictators do not know ex-ante how much the recipient has. The recipient asks for help by indicating the amount he requests from the dictator[2]. Dictators specify the amount they want to transfer for each possible request from the recipient; the computer then matches recipient's and dictator's choices.

The treatment manipulation is the ex-post information about the true endowment of the recipient dictators receive after they have made their decisions. In Treatment 'info', dictators learn the true endowment of the recipient at the end of the experiment. In Treatment 'no_info' dictators never learn the true endowment of the recipient. In Treatment 'not-to-know', dictators can choose (before sending the money to the recipient) if they want to receive ex-post information about the true endowment of the recipient.

First two treatments with exogenous ex-post disclosure manipulation were run within-subjects; the treatment with optional ignorance is run in separate sessions. Participants interacted over two periods with perfect stranger matching and no feedback between periods. I ran 6 experimental sessions in October 2014 in the laboratory of the Max Planck Institute for Economics in Jena. In total 188 participants took part in the experiment.

Main findings:

- Result 1: Dictators in 'not-to-know' treatment shared significantly less; no difference between treatments with and without exogenous ex-post information disclosure;

[2]Recipients could ask for any amount between 0 and 5 EUR or choose not to send any request at all.

- Result 2: Dictators do choose to remain ignorant about recipients' endowment. The frequency of 25%-46% 'not-to-know' choices shows that a significant minority of dictators choose to remain ex-post ignorant about recipient's endowment.
- Result 3: Those dictators who choose not to receive the information about recipient's endowment share significantly less than those who choose to reveal.

These findings show that the ex-post information disclosure plays an important role in participant's pro-social behavior. Similarly to psychologically disadvantageous ex-ante information (in the spirit of moral wiggle room experiments), the possibility to remain ex-post ignorant to recipient's endowment enhances selfishness and sustains higher inequality.

1.2.2 Current and future effects of delegation of lying

Chapter 3 of the thesis (based on a joint project with Prof. Oliver Kirchkamp) contributes to the studies on the link between pro-social and ethical behavior by investigating current and future effects of possibility to delegate lying.

We build our intuition on the distinction between direct and indirect wrongdoing in moral psychology and argue that the possibility to delegate might significantly reduce the scope of lying aversion.

We also test how the delegation of lies affects subsequent compensatory behavior, the tendency suggested by moral balancing literature. We expected that delegation of lying reduces the need to compensate one's wrongdoing which manifests itself in lower sharing in a dictator game.

Experimental Design In the experiment participants play a three-person sender-receiver game followed by a dictator game. In the sender-receiver game, participants interact in groups of three: two 'senders' (player 1 and player 2) and one 'receiver' (player 3). The computer randomly allocates a fixed prize for the senders (80 ECU) and the receiver (40 ECU) among five virtual boxes and assigns each group to one of the two experimental conditions.

We compare two conditions: "no conflict" and "conflict" (within-subjects). In the "no conflict" condition, the monetary prizes for the senders and the receiver are placed in the same box; in the "conflict" condition, the prizes for the senders are placed in one box, and the prize for the receiver is placed in a different box.

After learning the allocation of prizes to boxes, senders are asked to specify the number of the box they advise receiver to open: "Your prize is in Box x" and to indicate if they would like to send the message *themselves* or to *delegate* (the word 'delegate' is not used in the instructions). After both senders have made their decisions, one of them is selected randomly and this sender's decision is implemented. If the selected sender has chosen to delegate, the number of the box advised by the other sender is sent to the receiver; if the selected player has chosen 'myself', the number of the box selected by this sender is sent.

Receivers then receive the message from one of the senders and are asked to choose a box they want to open[3]. The decision of the receiver defines the payment of the participants in the game.

In the dictator game, senders are asked how much of their earnings they would like to share with a new receiver.

We expected that the frequency of delegation is higher in the

[3]Receivers know if the delegation takes place.

conflict condition (where there is a need to lie), that senders who tell the truth themselves are more likely to delegate than liars, that senders who lie share more in the dictator game to compensate their wrongdoing, and that those who delegate share less that senders who lie directly due to lower responsibility.

We ran 7 experimental sessions in Experimental Laboratory of Max Planck Institute of Economics in Jena in November-December 2013. In total 204 participants took part in the experiment.

Main findings:

- Result 1: The frequency of delegation in "Conflict" condition is significantly higher.
- Result 2: Among those senders who delegate more choose to tell the truth themselves.
- Result 3: Sharing in "Conflict" condition in higher.
- Result 4: Senders who delegate share more than senders who lie directly.

We conclude that delegation possibility might indeed decrease the scope of lying aversion. If senders who otherwise tell the truth choose to delegate, the incidence of lying under delegation can be higher. Contrary to our intuition, senders who delegate the choice and are objectively less responsible for lying share more in the dictator game. We suggest that the delegation possibility serve as a screening device: senders with higher psychological costs of lying delegate and then use the chance to compensate their wrongdoing, whereas direct liars have lower psychological costs of lying and feel no need to balance their past transgression. More direct analysis of

the link between the responsibility behind the transgressions and the desire to compensate for it offers an intriguing venue for future research.

1.2.3 Experimental analysis of the effects of third-party action on moral behavior

Chapter 4 (based on a joint project with Dr. Matthias Uhl) contributes to the discussion of pro-social and ethical behavior over time by investigating the effects of a third-party action on the people's tendency to compensate their past actions. There are two opposing hypotheses about the direction of the link between past and future behavior: according to *moral balancing*, past and future actions are negatively associated, according to *moral consistency*, past and future actions are positively associated. Research has identified several factors determining whether balancing or consistent behavior emerges: monetary costs of previous actions, level of conceptual abstractions (recent vs. distant acts), and the mind-set (outcome vs. rule-based) predisposition. In this project we investigate if the mind-set categorization is important in understanding the link between one's own act and the action of the third party. We expect that outcome-minded participant react to actions of others as to substitutes and reduce their subsequent generosity in case they face a good deed by others, whereas rule-based participants react to actions of others as complements and decrease their generosity once they face a bad deed by others. The experiment provided in this chapter tests these propositions.

Experimental Design The experiment consists of three parts: recall task, dictator game, and real-effort task. In the recall stage,

participants are instructed to recall either their recent ethical or their recent unethical action (between-subjects). Between the recall and the dictator game, participants are informed about the amount of efforts they will have to exert in the real-effort task at the end of the experiment. The amount of efforts can be 'low' (only a few tables to solve) or 'high' (a large number of tables to solve). In the baseline, the amount of efforts is defined by the computer. In the 'third party' treatment, the amount of efforts is defined by another participant. At the end of the experiment participants had to resolve a standard trolley dilemma: those who state it is appropriate to pull the leverage and save five people by sacrificing one person are treated as 'outcome-based' and those who state it is inappropriate to pull the leverage are treated as 'rule-based' participants.

We expected that recalling unethical action would prompt outcome-based participants to give more in a subsequent stage (exhibit compensatory behavior) and rule-based participants to give less (exhibit consistent behavior). We further explore how good or bad action from the third party affects these tendencies.

We ran 7 experimental sessions in August 2014 in Experimental Laboratory of Max Planck Institute of Economics in Jena. 218 participants in total took part in the experiment.

Main findings:

- Result 1: Our manipulation in the recall task induces consistent behavior. Participants in 'unethical' group are less generous than participants in 'ethical' group. The effect is more prominent for out-come based participants.

- Result 2: Outcome-based participants do not react to third-party action, but rule-based participants do. 'Unethical' rule-based participants share significantly less in the treatment.

We conclude that outcome-based vs. rule-based distinction is important to understand the dynamics of ethical behavior. We speculate that rule-based participants consider the actions of others as examples to follow and react asymmetrically stronger to unethical actions of others.

Chapter 2

The effect of ex-post information disclosure on giving in a dictator game with incomplete information

2.1 Motivation

People readily help others in need. However, requests for help might come not only from sincere recipients who have no money to cover expensive medical treatment of their children or are unable to repay debts due to sudden illness but also from dishonest recipients who simply try to exploit the credulity of donors. If there is no possibility of differentiation between the types of recipients, this might severely undermine the willingness to provide help by donors. In other words, dishonest applicants impose negative externalities on poor recipients and distort the willingness to donate. The reluctance to help in such environments with incomplete information might be largely driven by the psychological costs donors bear in case their help is misdirected. I employ a dictator game with incomplete information and manipulate the ex-post disclosure of the true need of the recipient to study donors' reaction towards donation

requests.

2.2 Related literature

Asking and giving Donors react to expectations of recipients and to the requests for help they receive. This tendency is observed in dictator-like experiments with communication (Yamamori *et al.*, 2008; Langenbach, 2014) and is often captured by belief-dependent models like guilt-aversion by Battigalli & Dufwenberg (2007)). The conclusion from these theoretical and empirical studies is that donors do not donate due to the dislike of inequality (Fehr & Schmidt, 1999; Ockenfels & Bolton, 2000) or pure altruism (warm-glow by Andreoni (1990)) but rather condition their giving on what the recipients expect and what they ask donors for.

Khalmetski *et al.* (2013) provide dictators with first order beliefs of recipients and find both positive and negative correlation between giving and expectations. Andreoni & Rao (2011) allow recipients to directly 'ask' a dictator for a transfer and observe that the communication of reasons and numerical requests substantially increase transfers. Interestingly, when only donors had a one-way communication possibility they decrease their giving compared to the baseline, arguably by justifying their greediness. Yamamori *et al.* (2008) equip recipients with the pay-off irrelevant minimal offer requests and observe intriguing heterogeneity in dictators's responses: lower requests trigger lower donations, and 'fair' requests (half of the pie size) often result in perfectly equal allocations; high ('greedy') requests, in turn, are sometimes punished by lower donations. These findings suggest that asking for a donation is an effective and sophisticated strategy for the recipients.

Informational asymmetry Another body of literature suggests that pro-social behavior is sensitive to various forms of informational asymmetry. On the one hand, informational advantage of dictators seem to dramatically decrease generosity (Gueth *et al.*, 1996; Andreoni & Bernheim, 2009; Dana *et al.*, 2007; Lazear *et al.*, 2012) ('dark side' of ignorance). On the other hand, informational disadvantage is shown to increase trust (Aimone & Houser, 2013) and enhance transfers to recipients under uncertain efficiency gains (Winschel & Zahn, 2014) ('positive side' of ignorance). Interestingly, the ambiguous effect of ignorance remains under the possibility for participants to completely resolve uncertainty. In one of the treatments, Dana *et al.* (2007) allow dictators to choose if they want to know the consequences of their actions for the recipient ('reveal the game'). Dana *et al.* observe that many dictators have chosen not to reveal the game and that the frequency of selfish choices increased dramatically under ignorance possibility. In contrast, in Aimone & Houser's trust game with the possibility for the trustors 'not-to-know' the action (the possible betrayal) of their trustee, the incidence of trust was considerably higher.

Ex-post information disclosure To our best knowledge, this is the first study of both asking and giving under informational asymmetry about recipient's need. In such settings, an inference about expectations of recipients and assessing the credibility of their requests becomes more difficult. Donors can anticipate emotional disutility from possibly helping the 'wrong' recipient as well as additional pleasure of helping those in need. If the psychological costs are greater than foregone benefits, donors might reduce their donations in such setting. The resulting unwillingness to donate that we expect resonates with lower trust due to anticipated costs of betrayal (Bohnet & Zeckhauser, 2004) or more general regret aversion

phenomenon documented for individual decision making under risk and uncertainty (Lovelady, 2014).

I suggest that experiencing these extra emotional costs/benefits begins at the moment of revelation of outcomes and intentions. If that is true, receiving no information about the true endowment of the recipient should prohibit emergence of this disutility and help to increase donations. At the same time, the possibility of ignorance might help to sustain self-serving belief that the recipient does not have a need.

I model the need as having zero earnings for a real-effort task with possibility of sending numerical requests to donors with higher endowment (Yamamori *et al.*, 2008). In contrast to previous studies of the effect of the *ex-ante* information about the consequences of their action for the affected party, I analyze dictator's reaction towards the *ex-post* disclosure of recipient's endowment.

2.3 Research questions

- Do recipients take advantage of informational asymmetry, i.e., do highly endowed recipients 'hide' behind the needy and demand larger donations?
- Do donors condition their giving on the requests, i.e, do donors 'trust' the requests from recipients?
- Do donors choose to remain ignorant about the true type of the recipient?
- How does (the choice for) ex-post disclosure of the true need affect dictator's reaction towards donation requests?

2.4 Experimental Design

I employ a modified dictator game with requests and incomplete information about recipient's endowment.

At the beginning of the experiment participants perform a real-effort task: they have to count 1s in a 10 by 10 table and solve eight such tables correctly. For completion of this task half of the participants receive a sure payment of 10 EUR (dictators), the earnings of the other half of the participants (recipients) depend on the individual random draw: a participant receives either low (0 EUR) or high (6 EUR) endowment with equal probability. The endowment of the recipients is private information. After the real-effort task each Dictator is matched with one Recipient.

The recipient communicates to the dictator how much he would like to receive from him. Recipients can choose to request an amount X between 0 and 5 EUR (in steps of 50 cents) or choose not to send any request at all. Dictators choose an amount Y between 0 and 10 EUR (in steps of 10 cents) they would like to send to the recipient without knowing the exact endowment of the recipient. Dictators' choices are elicited via strategy method: they specify the amount of the transfer for each possible request from the recipient; the computer then matches the corresponding recipient's and dictator's choices and transfers the respective amount from the dictator to the recipient.

As a treatment manipulation I change the ex-post information dictators receive about the true endowment of the recipient. In Treatment 'info', dictators learn the true endowment of the recipient at the end of the experiment. In Treatment 'no_info', dictators never learn the true endowment of the recipient. In Treatment 'not-to-know', dictators can choose (before sending the money to

	Recipient's endowment	**Ex-post disclosure**
Treatment "info"	0 or 6 EUR	Yes
Treatment "no_info"	0 or 6 EUR	No
Treatment "not-to-know"	0 or 6 EUR	Optional

Table 2.1: Experimental Conditions

the recipient) if they want to receive ex-post information about the true endowment of the recipient. The recipients are not informed if the dictators use this opportunity.

Table 2.1 summarizes the treatments of the experiment.

During the experiment participants' earnings was expressed in Experimental Currency Units (ECU) and converted into EUR at the rate 10:1 at the end of the experiment.

For the first two treatments I employ a within-subject design. Participants interact over two periods with perfect stranger matching and no feedback between the periods.

We vary the order of the conditions: half of the groups are assigned to "no_info" → "info" sequence; the other half is assigned to "info" → "no_info" condition in the 1st and the 2nd period respectively.

Participants are told that one of the two periods will be randomly chosen for payment and that there is 1/2 probability that the earnings of the recipient will be disclosed.[1]

At the beginning of the period dictators are informed whether they will receive the information about the earnings of the recipient

[1]Neither dictators nor the recipients know upfront that the information condition varies across periods.

(once this period is selected for payment); recipients do not receive this information. [2]

The experiment concludes with the empathy questionnaire.

The treatment 'not-to-know' is run with separate sessions. The instructions are identical to the other two treatments apart from the possibility of ex-post information acquisition by the dictator. In each period, before the dictator decides about the transfer, he is asked if he wants (at the end of 2nd period) to disclose the information about the earnings of the recipient (for each possible request). The recipient is not informed about the choice of the dictator to reveal his earnings (see instructions).

At the end of 2nd period we elicit unincetivized beliefs of dictators about the average requests from recipients (dictators are asked to guess the average requests by low and high endowed recipients); recipients are asked to guess the average transfer from dictators.

2.5 Hypotheses

I expect that incomplete information setting creates negative externalities and diminishes the possibility of requests to be met by donors:

Hypothesis 2.1. *Dictators 'reject' large requests from recipients; resulting allocations are unequal*

If donors reject large requests and treat them as not informative about recipient's real need, poor recipients have no chance to improve their well-being by asking for more.

[2]Since the recipients do not learn the condition they are assigned in a particular period, this restriction keeps the expected probability of their type to be ex-post discovered constant across treatments.

Hypothesis 2.2. *Ex-post information disclosure affects giving of the dictators*

Receiving no ex-post information about the true endowment of the recipient should disable experiencing psychological costs of helping non-deserving recipients. Here I take an optimistic position and hope that positive side of less information in such an environment is stronger than the dark side of possible ignorance. Since the real-effort task should create the moral obligation to help the poor recipient, we believe that ignoring the requests in our setting is more difficult than simply ignoring the consequences for others in the spirit of moral wiggle room experiments.

If ex-post information hurts both selfish and pro-social dictators, the costless possibility of ignorance should be used:

Hypothesis 2.3. *Significant fraction of dictators choose "not-to-know" the information about the true endowment of the recipient.*

The willingness to disclose the earnings of the recipient might depend on the amount requested. Dictators might be more willing to reveal recipients' endowment for large rather than for small requests. If the "not-to-know" option indeed allows dictators to avoid psychological costs of being exploited, one should observe that:

Hypothesis 2.4. *Giving to low endowed recipients in "not-to-know" Treatment is higher compared to the full information condition.*

If, however, the "not-to-know" option is predominantly used as a self-serving ignorance device, one should observe that not-to-know option does not increase the transfers to low endowed recipients.

2.6 Results

I ran 6 experimental sessions in October 2014 in the laboratory of Max Planck Institute for Economics in Jena. The experiment has been implemented in z-Tree (Fischbacher, 2007), participants have been recruited with ORSEE (Greiner, 2004). 188 participants (94 recipients and 94 dictators) took part in the experiment. The average age of the participants was 23.4; gender composition was: 59% females and 41% males.

Giving by Treatment

I observed relatively low willingness to share money: on average in all of the conditions participants in the role of dictators transferred less than 8 ECU (8% of their endowment) to recipients. At the same time, only a minority of dictators behaved purely selfishly and transferred nothing for each possible request from the recipient: 23% and 29% in Period 1 and 2 respectively.

Figure 2.1 summarizes the average amounts transferred in each of the conditions.

As Figure 2.1 shows, dictators condition their giving on recipient's requests. The transfers increase with the amount requested for moderate requests up to 20 ECU and tend to decline for the requests above 20 ECU [3]. The ex-post information disclosure does not affect aggregate sharing for requests up to 20 ECU; dictators in "not-to-know" treatment give less than in the other two treatments for the range of requests higher than 20 ECU. These findings resonate with our Hypothesis 2.1: dictators do not increase their

[3]The transfer of 20 ECU is pay-off equalizing donation in case the recipient has a high endowment of 60 ECU; dictators rarely give more than 20 ECU.

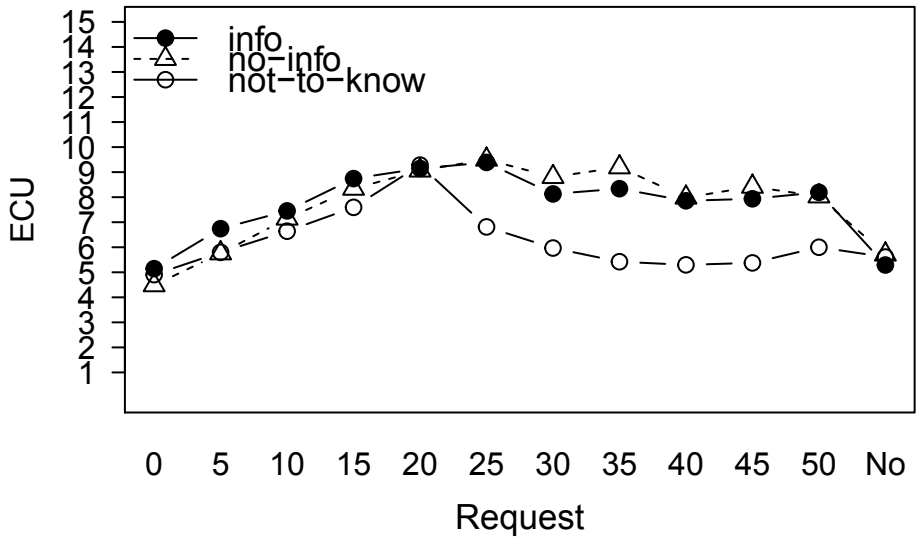


Figure 2.1: Average transfer by request and condition, ECU

transfers with the larger requests, recipients with low endowment 'suffer' from incomplete information setting.

To better understand the differences in donations across treatments we estimate the following mixed-effects model:

$$\begin{aligned} \text{Give} = \beta_0 + \beta_1 d_{\text{Request\_high}} + \beta_2 d_{\text{Condition}} + \\ \beta_3 d_{\text{Request\_high}*\text{Condition}} + \gamma_t + \epsilon_i + \epsilon_g + \epsilon_{igt} \end{aligned} \tag{2.1}$$

$d_{\text{Request\_high}}$ is a dummy which is one when the request exceeds 20 ECU,$d_{\text{Condition}}$ is a dummy for treatment,γ_t is a fixed effect for period t, ϵ_i is a random effect for the individual, ϵ_g is a random effect for the matching group, and ϵ_{igt} is the residual.

Table 2.2 summarizes the estimation results.

For all regressions in Table 2.2 we use the amount of ECU participants share as the dependent variable.

Table 2.2: Mixed-effect regression estimates: DV - amount shared, ECU

	all dictators	no selfish dictators	all dictator:order
Intercept	7.22***	8.71***	7.81***
	(1.15)	(1.17)	(1.51)
request_high	1.22**	1.51**	1.22**
	(0.45)	(0.59)	(0.45)
no_info	−0.34	−0.35	−0.34
	(0.45)	(0.60)	(0.45)
not_to_know	−0.45	1.39	−1.04
	(1.94)	(2.04)	(2.17)
Period_2	−0.28	−0.05	−0.28
	(0.26)	(0.36)	(0.26)
no_info*request_high	0.69	0.95	0.69
	(0.64)	(0.84)	(0.64)
not_to_know*request_high	−2.05**	−2.80**	−2.05**
	(0.64)	(0.87)	(0.64)
Info_1st			−1.18
			(1.96)
Num. obs.	2256	1668	2256
Num. groups: i	94	73	94
Num. groups: g	6	6	6

***$p < 0.001$, **$p < 0.01$, *$p < 0.05$

First column in Table 2.2 'all dictators' includes all participants in the role of dictators, 2nd column 'no selfish dictators' excludes those participants who shared nothing for each request of the recipient, and the third column 'all dictators:order' includes an order dummy variable 'Info_1st' (takes value of one if subjects started with 'info' condition in the first round of the experiment).

As Table 2.2 shows, large requests (dummy variable 'request_high') on average trigger larger donations. Ex-post information condition itself (variables 'info' and 'not-to-know') does not significantly affect

the rate of aggregate giving. However, for large requests (requests above 20 ECU) participants in 'not-to-know' condition shared significantly less. Therefore, our Hypothesis 2.2 about the relevance of ex-post information disclosure is partially corroborated.

This observation suggests that dictators might try to 'ignore' large requests by choosing not to receive information about the true endowment of the recipients.

This brings us to our next hypothesis.

Information acquisition

Do dictators choose not to receive ex-post information?

Figure 2.2 shows the faction of not-to-know choices in each of the two periods.

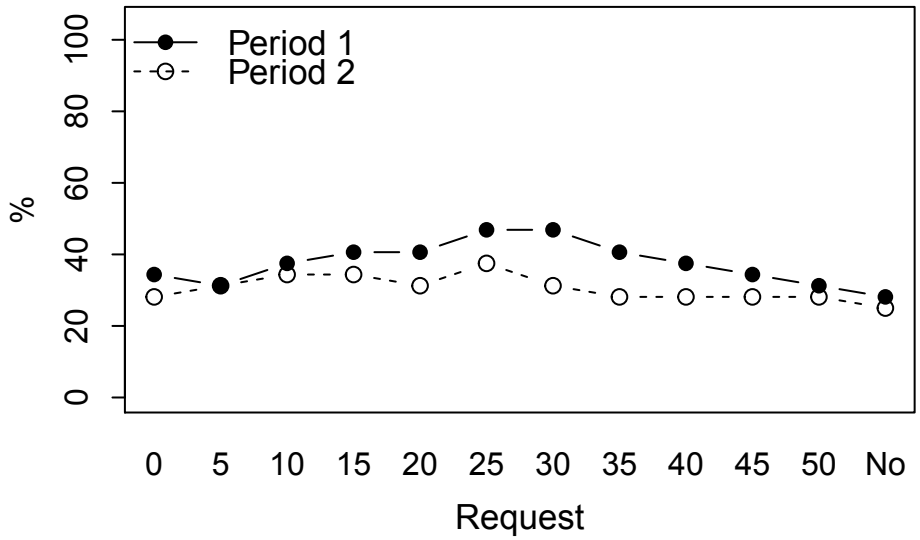


Figure 2.2: Fraction of "not-to-know" choices by request, %

As Figure 2.2 shows, substantial minority of dictators' (from 25% to 46%) choices is not to reveal the information about the true endowment of the recipient. This supports our Hypothesis 2.3: many dictators prefer to remain ignorant about recipients' endowment. Interestingly, the number of 'not-to-know' choices peaks at moderate requests (25-30 ECU), requests which can potentially call for fair splits. Although the majority of dictators do not condition their choices of information disclosure on the amount requested, significant minority of subjects 28% (Period 1) to 25% (Period 2) switch their decision [4].

Figure 2.3 depicts average transfers from dictators in "not-to-know" Treatment.

As Figure 2.3 shows, dictators who choose to (ex-post) reveal the information about the endowment of the recipient transfer more money to the recipients than dictators who prefer to stay ignorant.

To test if the dictators who choose not to reveal share significantly less, I compute individual average amounts transferred across the requests. I find that dictators who (more frequently) choose not to reveal recipient's endowment shared significantly less compared to dictators who (more frequently) choose to disclose (p.value of Wilcoxon rank sum test is 0.015)[5]. Notably, the fraction of selfish dictators (transferring zero to the recipient in the first period) is three times as large among dictators who choose remain ignorant (53% vs. 17%, p.value of Fisher exact test is 0.056)[6].

[4]Switching behavior is heterogeneous: some participants start with 'not-to-know' for low requests and then want to disclose the endowment of the recipients who ask for more than 20 ECU, the others choose to reveal for low requests and then prefer to keep their eyes closed.

[5]For the purpose of statistical comparison I classify dictators by the majority of their 'not-to-know' choices: if a dictator chose to reveal recipient's endowment for more than a half of the requests, I classify him as willing to disclose recipient's endowment and as preferring not to disclose the recipient's endowment otherwise.

[6]All subjects (6 out of 32 dictators) who chose to ignore for each possible request shared

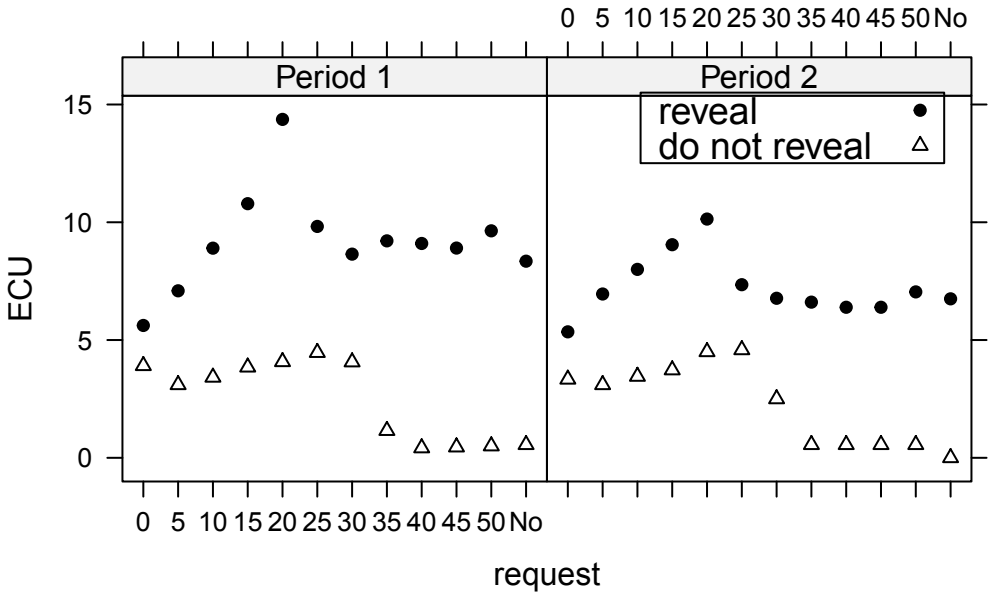


Figure 2.3: Average transfers by not-to-know choice, ECU

Truning to Hypothesis 2.4, the possibility of ignorance indeed affects giving, namely it increases inequality in incomplete information setting. I do not find evidence of positive side of ignorance in the experiment.

These findings suggest that not only possibility of no ex-ante information disclosure but also ignorance towards ex-post information might inhibit pro-social behavior.

nothing with the recipient.

2.7 Conclusion and Discussion

I analyzed experimentally the effect of ex-post information disclosure on the dictator's giving under incomplete information where dictators face uncertainty about the endowment of the recipient and the recipients communicate their requests for the desired donation. In line with previous experiments, I observe the link between asking and giving: dictators condition their transfers on how much recipients ask. As regards different information conditions, I observed no significant effect of exogenous ex-post information manipulation: giving in 'no-info' and 'info' treatments is not significantly different from each other. In both treatments dictators increase their giving for the range of modest requests and decrease their giving for too ambitious requests.

In the treatment where dictators could choose not to reveal the ex-post information, the average giving is considerably lower. The participants who choose not to reveal the information give less that the participants who want to see the consequences of their transfer.

The analysis of participants reasoning confirms the relevance of ex-post information. Many dictators report they would like to balance between the profit and their 'bad conscience'. Some of dictators indicate they 'hoped' that their recipient had higher endowment and therefore transferred small amounts. However, only a few dictators indicated directly that they made use of ignorance possibility: 'Today I wanted to receive the maximum pay-off and have decided to give nothing to participant B. I did not want to see much he asked for so that I do have a bad conscience in case he asked a little', 'In the first part of the experiment I decided to share as long as I receive the information about the profit of participant B... In the 2nd part, however, there was no feedback. That is

why I decided to give nothing'. The fact that only in 'not-to-know' treatment dictators are explicitly asked to choose if they want to get informed might have made this decision more salient compared to 'no_info' treatment. In exogenous information conditions, dictators seem to be more focused on the reaction to recipients request (to punish 'the greed') rather than on the ex-post feedback about recipient's endowment.

I conclude that the ex-post information disclosure plays an important role in participant's pro-social behavior. Similarly to psychologically disadvantageous ex-ante information (in the spirit of moral wiggle room experiments), the possibility to remain ex-post ignorant to recipient's endowment enhances selfishness and sustains higher inequality.

2.8 Appendix: Experimental Instructions

*(*Common text for all treatments. Text specific for 'not-to-know' treatment is written separately).

Welcome and thank you for participating in the experiment of the Max Planck Institute for Economics. Please switch off your mobile phones and remain silent. All the decisions you make in this experiment will be treated anonymously. Any communication between the participants is not allowed. Should you have any questions during the experiment, please raise your hand, and one of the experimenters will answer you privately.

For your participation in the experiment you will receive 2.5 EUR show-up fee. In the course of the experiment you can earn additional money. Your earnings in this experiment depend on the decisions you make, the decisions of the other participants as well as on chance. During the experiment your earnings will be expressed in ECU (Experimental Currency Units) which will be translated into EUR at the end of the experiment at the following exchange rate:

10 ECU = 1 EUR.

The course of the experiment

In the first stage of the experiment all the participants will be asked to perform the following task: you will be presented a series of 10 by 10 tables containing 0s and 1s in a randomized order. The example of such a table is presented below:

0	0	1	1	0	1	1	0	0	1
1	1	1	0	1	1	1	1	1	0
1	1	1	0	1	1	0	0	1	1
0	0	1	0	0	0	1	0	0	1
1	1	0	1	0	1	1	1	1	0
0	0	1	0	1	0	1	0	1	1
0	1	1	0	1	0	0	0	0	0
0	0	1	1	1	0	1	0	1	1
1	1	0	0	0	0	0	1	0	0
0	0	1	1	1	1	1	0	1	1

When you see a table, your task is to compute the number of 1s in that table. Each participant will have to correctly solve **eight** such tables. If the number of 1s you enter is wrong, you will receive a new table until you accumulate 8 correct answers (you cannot go back to the same table again).

After completion of the task half of the participants would be randomly assigned to one of the two roles: **Participant A** and **Participant B.**

Your earnings of the participants for this task in defined as follows:

Participants A will receive a sure payment of **100 ECU** for the task;
Each participant B will receive either **60 ECU** or **0 ECU** for the task with equal probability. (The random draw defining whether the participant receives 0 ECU or 60 ECU for the task is exercised by the computer for each participant B individually).

After the computer defines the payments, the participants will be informed about their earnings on their screens.

The experiment will then proceed over two periods.

At the beginning of each period, the participants will be paired: each Participant A will be matched with one Participant B. Matching procedure guarantees that in the next period no participant will be matched with the same participant he interacted in the first period of the experiment.

Participant B will then have the opportunity to communicate to Participant A an **amount X** between 0 and 50 ECU he would like to receive from him. Participant B can ask for 0;5;10;15;..;50 ECU or choose not send any request at all.

Participant A, without knowing if Participant B has earned 0 or 60 ECU for the task, will have the opportunity to transfer an **amount Y** between 0 and 100 ECU to Participant B. Before Participant A knows the exact amount requested by Participant B (or whether Participant B chose not to send any request) he will be asked to specify the amount he would like to transfer for each possible request from Participant B. The computer will then match the amount requested by Participant B and the transfer specified by Participant A for that case. The respective transfer specified by Participant A will be subtracted from the earnings of Participant A and added to the earnings of Participant B. (The other transfers specified by Participant A for the other possible requests not chosen by Participant B will have no effect on the earnings). The earnings of the participants:

Participant A: **100 - Y ECU**;
Participant B: **0 + Y ECU** or **60 + Y ECU** respectively.

After the two periods are over, the computer will randomly select

one of the two periods for payment. Participants will be informed about the amount transferred by Participant A, the request of Participant B and (with 1/2 probability)[7] the earnings of Participant B.

Since both periods have the same chance to be chosen for payment, you have to decide carefully in each of the two periods.

The experiment will proceed to a questionnaire.
If you read the instructions carefully and do not have further questions, please click on "Continue".

[7]In 'not-to-know' treatment instead of 'with 1/2 probability' participants read 'if requested by Participant A'

Decision Screen for Recipients:

Period 1.
Your role is Participant B.
Your endowment is: 0 (60) ECU.
Endowment of Participant A: 100 ECU.

Please make you decision about the amount of ECU you would like to request from Participant A.

	0	5	10	...	30	35	40	45	50	no request
Your choice:				...						

Decision Screen for Dictators('info' and 'no_info' treatments):

Your role is Participant A.
Your endowment is: 100 ECU.
Endowment of Participant B: 0 or 60 ECU

If this period is chosen for payment, you will (NOT) be informed about the earnings of Participant B.

Please make you decision about the amount of ECU you would like to transfer to Participant B.

Participant's B request:	0	5	10	...	30	35	40	45	50	no request
Your transfer:				...						

Decision Screen for Dictators in 'not-to-know' treatments:

Please decide if you want **at the end of the two periods** to receive the information about the earnings of Participant B :

Participant's B request:	0	5	10	...	30	35	40	45	50	no request
Information disclosure	Yes No	Yes No	Yes No	... ...	Yes No	Yes No	Yes No	Yes No	Yes No	Yes No

Participant B will not be informed whether you have chosen to disclosure his earnings.

Please make you decision about the amount of ECU you would like to transfer to Participant B.

Participant's B request:	0	5	10	...	30	35	40	45	50	no request
Your transfer:				...						

Chapter 3

Do I care if the others lie? Current and future effects of delegation of lying

3.1 Introduction

3.1.1 Motivation

Lying brings benefits but comes at a cost. Reluctance to benefit from sending untruthful information in an anonymous setting, known as *lying aversion*, is of continuing interest in behavioral economics. Many empirical and theoretical studies of the unwillingness to lie to the disadvantage of others (Gneezy, 2005; Fischbacher & Föllmi-Heusi, 2013; Erat & Gneezy, 2012; Kartik, 2009) want to answer the question why people are telling the truth: is it a desire to respect trust, to avoid losses for others, or a mere distaste for lying per se?

Baron & Ritov (2004); Spranca *et al.* (1991); Royzman & Baron (2002) distinguish direct and indirect negative consequences of actions. They find that negative outcomes are perceived as more

harmful if they are the result of a direct action and as less harmful if they are the result of an indirect action. We suspect a similar distinction between the consequences of direct and indirect lies. People might prefer not to lie. However, they might consider an indirect lie, i.e. a lie through an intermediary, as more acceptable than an own lie. As a reason we suggest that people not only want to avoid the harm or the violation of trust but also want to avoid the directness of the wrongdoing.

Hamman *et al.* (2010); Bartling & Fischbacher (2012); Coffman (2011) find that delegation reduces responsibility and makes it easier to reach self-interested or immoral allocations. Still, people might view delegation differently ex-ante and ex-post. In line with Nisan & Horenczyk (1990); Sachdeva *et al.* (2009); Gneezy *et al.* (2012a) we suspect that delegation could influence people's ex-post compensatory behavior. After delegation people might be more ready to cleanse their past wrongdoing.

We study a game where it is possible to delegate the act of lying and where it is possible to take pro-social actions subsequently. We examine how delegation affects the outcomes of people's current and future ethical decisions.

3.1.2 Related literature

Cause and effect of delegation in the positive and the negative domain Efficiency could be a standard reason to delegate: an agent could be better equipped with resources, time, or expertise. A different motive has been brought forward by Hamman *et al.* (2010): people who are reluctant to implement painful decisions (selfish allocations, discriminatory judgments, outright lies) themselves might find that delegation helps them to avoid the disutility from a direct

harmful act and free them to act in their best interest. Hamman *et al.* compare a standard dictator game with and without the possibility of delegation. Without delegation they find a substantial fraction of fair allocations. When delegation is possible, many principals delegate to agents, and those agents, who are not directly incentivized for a principal-favouring allocation, act more in the interest of their principals than the principals themselves. As a result, delegation substantially increases inequality. Hamman *et al.* suggest that shifting (and diffusion of) responsibility explain their result: principals and agents share and thereby reduce the joint responsibility for their actions.

Bartling & Fischbacher (2012) use delegation as a workhorse to compare different reasons for punishment: outcome, intention and responsibility. They observe that delegation reduces punishment. Furthermore, responsibility has a larger impact on punishment than outcome and intention. In line with Hamman *et al.*, delegation shifts and dilutes responsibility.

Coffman (2011) considers a game which allows to distinguish two causes for punishment: responsibility and directness of the interaction. In Coffman's game responsibility is not affected by intermediation. Still (and in line with Hamman *et al.*, 2010; Bartling & Fischbacher, 2012) Coffman observes that intermediation reduces punishment. Coffman can further say that it is the indirect interaction, not the shifting of responsibility, which leads a reduction in punishment.

How intermediation affects the moral cost of a transaction is studied by Drugov *et al.* (2014) who use a bribery game. Drugov *et al.* find that intermediaries facilitate corruption not by reducing the responsibility for the outcome but rather by replacing a direct with an indirect link.

The concept of "moral distance" from a negative outcome mentioned by Drugov *et al.* is long known in moral psychology. Also the detrimental effects of indirect action are well-documented for both self- and other-regarding decisions (Baron & Ritov, 2004; Spranca *et al.*, 1991; Royzman & Baron, 2002; Hayashi, 2013). Inasmuch as dictators are held less responsible if they delegate (Hamman *et al.*, 2010; Bartling & Fischbacher, 2012; Coffman, 2011; Drugov *et al.*, 2014), allocations by *omission* trigger less blame by the recipients (DeScioli *et al.*, 2011).

But not only for morally questionable actions, also for desirable actions we find a distinction between the direct and the indirect. While decision makers prefer to implement unethical actions indirectly, i.e. through an intermediary, the same decision makers prefer to implement benevolent activities (generous donations, non-discriminatory judgments, honesty) rather directly. Patt & Zeckhauser (2000), for example, model willingness to attribute positive outcomes to one's own actions and provide the evidence of 'action bias' in environmental decisions: people prefer actively implementing environmentally friendly policies even though inaction would lead to better environmental outcomes. Coffman (2011) compares direct (donor-recipient) and indirect (donor-fund-recipient) donations and finds that people reward donors much less if they donate to a cause through an intermediary. Eisenkopf & Fischbacher (2011) investigate the same reward pattern in a trust game. In their setting with two trustors and one trustee, delegation by the first trustor to the second one can potentially increase efficiency. They find that trustees seem not to recognize that the efficiency gain is due to delegation by the first trustor. Trustees do not reward the first trustor correspondingly.

Lying aversion and delegation People do not always lie even if lying secures high monetary reward.

In a seminal experiment, Gneezy (2005) employs a deception-game to test for (non-strategic) *lying aversion* — reluctance to get the outcome through lying. In his setting, a sender learns about the distributions of pay-offs behind option A and option B and advises a receiver which of the two options to choose: 'Option A (B) will earn you more than Option B (A)'. Since the senders' payoffs are high when receivers' payoffs are low and vice versa, and since receivers do not know this, senders have an interest to lie. Gneezy compares choices in deception and in dictator games with equivalent pay-offs and finds that the fraction of selfish choices in dictator games to be higher than the fraction of lies in deception games. Gneezy concludes that lying is not neutral.

Since then a number of studies on various aspects of lying aversion have appeared (see Erat & Gneezy, 2012; Vanberg, 2008; Fischbacher & Föllmi-Heusi, 2013; Sutter, 2009). It has been shown that the expectations of the receiver, the damage from lying, and the observability of lies shape but do not fully explain preferences for truth-telling.

Although different motivations for lying aversion have been addressed, it remains unclear whether people distinguish between direct (own) lies and indirect lies (lies by an intermediary).

To shed light on this issue Erat (2013) studies a three-person sender-receiver game where senders can delegate. Receivers are not aware of the alignment of interests and receivers do not know who of the two players sends the message. Erat observes that roughly 30% of senders delegate the decision. Erat also finds that an increase in the receiver's cost of deception does not increase truth-telling but does increase delegation.

Compensatory behavior and lying Since senders even delegate when receivers do not know who sent a message, this suggests that one motive for delegation might be the preservation of the self image. In this case deception could also be linked to subsequent compensatory behavior of the sender.

According to moral balancing theories (Nisan & Horenczyk, 1990; Meritt *et al.*, 2010; Sachdeva *et al.*, 2009) people form a (subjective) benchmark of acceptable morale and allow for both positive and negative deviations as long as balance is appropriate. In this framework, doing extra good (creating a surplus to the moral account) may 'license' a subsequent bad action, and doing extra bad (creating a moral debt) may be 'cleansed' or 'compensated' by a future good deed to restore the account.

Moral cleansing, the desire to compensate a bad action with a following good act, is sometimes explained within self-signaling models (Benabou & Tirole, 2011), where individuals with no perfect access to their deep preferences might 'invest' in a bad behavior to get a signal of their true (good) type prompting higher goodness in the subsequent task. This reasoning is in line with the approach to underestimated emotions in economic behavior (Loewenstein, 2000), where moral cleansing may result from underestimation of the future negative emotions (like guilt or shame). If regret after lying is higher than expected, the initial choice turns out to be ex-post suboptimal which requires a compensation.

Gneezy *et al.* (2012b) discuss how feeling of guilt urges transgressors of the norm to behave more pro-socially. In Gneezy *et al.*'s experiment subjects who cheated in a first task contributed more to a charity than truth-tellers. Gneezy *et al.* conclude that an unannounced opportunity for pro-social behavior right after a transgression may serve as a conscience cleansing instrument and advises this

technique for charitable fund-raising.

If people exercise moral balancing, seek to cleanse a transgression, but account indirect harm differently, the possibility to delegate lying may decrease the positive compensatory behavior in the subsequent task.

In this project we want to examine the effects of delegation in a dynamic setting: first, we extend Erat (2013) and study how delegation affects intensity of lying (current effect); second, we investigate how delegation of lying affects subsequent compensatory behavior (future effect).

3.2 Experimental design

Our experiment lasts for four periods. In each period participants play a three-person sender-receiver game followed by a dictator game. In period 1 participants did not know about the upcoming dictator game; in periods 2 to 4 they knew about it. The experiment has been implemented in z-Tree (Fischbacher, 2007), participants have been recruited with ORSEE (Greiner, 2004). During the experiment payoffs are described as ECU. At the end of the experiment one period is chosen for payment and ECUs are converted into Euros at a rate of 10:1.

In the sender-receiver game, participants interact in groups of three: two 'senders' (player 1 and player 2) and one 'receiver' (player 3). The computer randomly allocates a fixed prize for the senders (80 ECU) and the receiver (40 ECU) among five virtual boxes and assigns each group to one of the two experimental conditions.

We compare two conditions: "no conflict" and "conflict". In the "no conflict" condition, the monetary prizes for the senders and the receiver are placed in the same box; in the "conflict" condition, the prizes for the senders are placed in one box, and the prize for the receiver is placed in a different box.

Figure 3.1 provides the examples of possible allocation of prizes in the two conditions (Prize i denotes the prize for player i).

Example for an allocation in the "no conflict" condition:

Box 1	Box 2	Box 3	Box 4	Box 5
Prize 1: 0	Prize 1: 0	Prize 1: 80	Prize 1: 0	Prize 1: 0
Prize 2: 0	Prize 2: 0	Prize 2: 80	Prize 2: 0	Prize 2: 0
Prize 3: 0	Prize 3: 0	Prize 3: 40	Prize 3: 0	Prize 3: 0

Example for an allocation in the "conflict" condition:

Box 1	Box 2	Box 3	Box 4	Box 5
Prize 1: 0	Prize 1: 0	Prize 1: 80	Prize 1: 0	Prize 1: 0
Prize 2: 0	Prize 2: 0	Prize 2: 80	Prize 2: 0	Prize 2: 0
Prize 3: 0	Prize 3: 0	Prize 3: 0	Prize 3: 0	Prize 3: 40

Figure 3.1: Information provided to senders in the two conditions

All players are informed about the two conditions, but only the two senders know the exact allocation of prizes to boxes; receivers are told they will be assigned to one of the two conditions with equal probability and will learn the exact allocation only after they make their decisions (See Figure 3.2).

After learning the allocation of prizes to boxes, senders are asked to specify the number of the box they advise receiver to open: "Your prize is in Box x". Together with the number of box they advise, senders are asked to indicate if they would like to send the message *themselves* or to *delegate* (the word 'delegate' is not used in the instructions). After both senders have made their decisions one of them is selected randomly (we will call this an "effective" sender

Box 1	Box 2	Box 3	Box 4	Box 5
Prize 1: ?	Prize 1: ?	Prize 1: ?	Prize 1: ?	Prize 1: ?
Prize 2: ?	Prize 2: ?	Prize 2: ?	Prize 2: ?	Prize 2: ?
Prize 3: ?	Prize 3: ?	Prize 3: ?	Prize 3: ?	Prize 3: ?

Figure 3.2: Information provided for receivers in both conditions

later) and this sender's decision is implemented: if the selected sender has chosen to delegate, the number of the box advised by the other sender is sent to the receiver; if the selected player has chosen 'myself', the number of the box selected by this sender is sent.

The receiver then obtains a message stating 'Your prize is in Box x' either from sender 1 or from sender 2, chooses a box and, thus, determines the pay-offs for all the players in the group. In contrast to Erat (2013), receivers know if the delegation takes place and learn if they were lied to in the feedback.

In the subsequent dictator game, participants keep their roles, but groups are re-matched. Each participant meets two other participants he or she did not interact with in the sender-receiver game. Participants in the role of senders do not know what the new receiver earned in part 1. Senders decide how much out of their 80 ECU earnings they would like to transfer to the new receiver. The computer then randomly selects one of the two senders and implements his decision.

Participants repeated this interaction (sender-receiver + dictator game) for four periods (random matching). Each sender played twice in each of the two conditions in the following order: C-C-NC-NC (half of the groups) or NC-NC-C-C (half of the groups), where C stands for "conflict" and NC stands for "no conflict" condition; receivers were randomly assigned to one of the two conditions in

each period.

3.3 Hypotheses

3.3.1 Hypotheses for the sender-receiver game:

Since psychological costs are more likely to arise in the "conflict" and not in the "no conflict" condition, we expect more delegation in the "conflict" condition as a means to reduce psychological costs behind lying.

Hypothesis 3.1. *The frequency of delegation is higher in "conflict" than in "no conflict".*

Let us now look at the differences in delegation behaviour of senders who lie and senders who tell the truth. If a sender who would otherwise lie delegates this sender achieves two things: the own responsibility and the related psychological cost is reduced but, since only some delegates will lie, also the expected payoff is reduced. For a sender who would otherwise tell the truth delegation also means a lower responsibility but, since at least some delegates will lie, the expected payoff increases. Delegation would then more more attractive for senders who tell otherwise the truth than it would be for senders who otherwise lie.

Hypothesis 3.2. *In the "conflict" condition, senders who tell the truth are more likely to delegate than senders who lie.*

3.3.2 Hypotheses for the dictator game:

Truth telling implies no norm violation (but no extra good to license negative behavior) and thus does not call for moral cleansing. In

turn, lying is a norm violation and induces negative feelings calling for cleansing. If our manipulation works and the majority of participants lie in the "conflict" condition, we should observe different shared amounts in the dictator game. This hypothesis is consistent with the observation of the lower donations by truth-tellers in Gneezy *et al.* (2012b).

Hypothesis 3.3. *Senders who have lied in the sender-receiver game share more in the dictator game than truth tellers.*

Since we expect more lying in the "conflict" situation, we also hypothesize the following:

Hypothesis 3.4. *Senders share more in the "conflict" than in the "no conflict" condition.*

Since all players get feedback who actually sent the message, senders can assess their own responsibility and can distinguish between indirect and direct outcomes. We expect that compensatory behavior is sensitive not only to outcomes but also to procedures. If, as we expect, indirect lying has a lower psychological cost than direct lying, we will see less cleansing behavior in the case of indirect decisions.

Hypothesis 3.5. *Conditional on the outcome of the sender-receiver game, senders who delegate share less than senders who lie directly.*

3.4 Results

We ran 7 sessions at the laboratory of the Max Planck Institute of Economics in Jena in November-December 2013. In total 204

subjects participated in the experiment. Sessions lasted for approximately 50 minutes. The average payment (including show-up fee) was 7.47 EUR.

3.4.1 Lying and truth-telling

The frequency of truth-telling in the two conditions is presented in Figure 3.3.

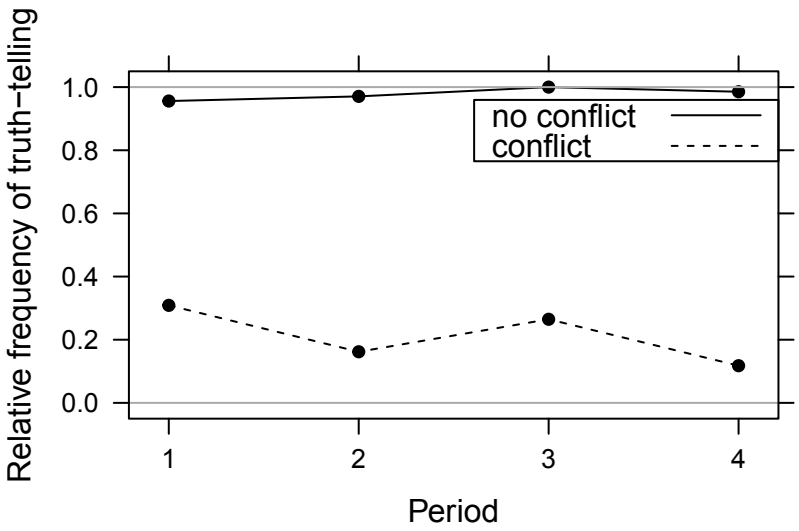


Figure 3.3: Truth-telling by Condition, n=68 in each period

Figure 3.3 shows the relative frequency of truth telling in the four periods of the experiment. We see a clear treatment effect: our manipulation had a desired effect on the level of truth-telling. In the "no conflict" condition, almost all (98%) participants send truthful messages. In the "conflict" condition only 21% messages are truthful. We should also note that even in "conflict" not all senders are lying. Consistent with the literature on lying aversion,

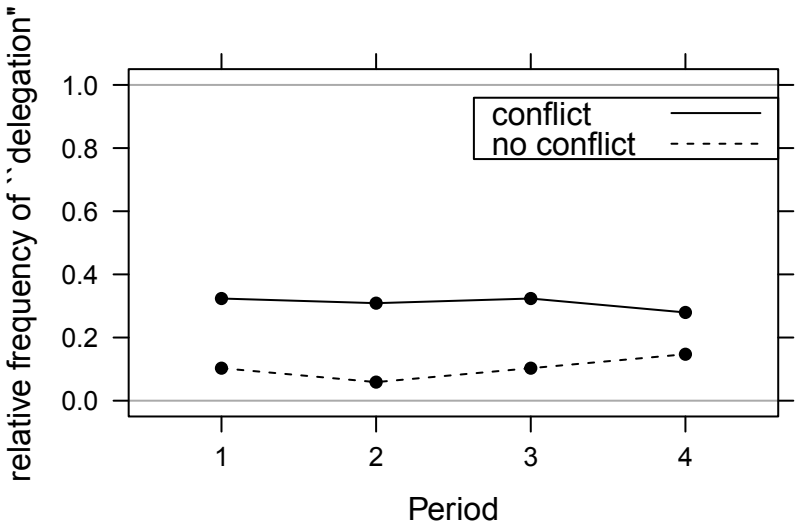


Figure 3.4: Delegation in "conflict" and "no conflict"

we observe 21% still tell the truth[1].

3.4.2 Delegation

Turning to Hypothesis 3.1 Figure 3.4 shows the fraction of senders who delegated their decision.

31% of all senders in "conflict" and only 10% of all senders in "no conflict" delegate. For the first period, where we can assume independence of choices, we apply a two-sample test for equality of proportions and find a significant difference between "conflict" and "no conflict" (p-value 0.0034). We, thus, confirm Hypothesis 3.1.

We next turn to Hypothesis 3.2. Figure 3.5 shows fractions of del-

[1] Although one can argue that some of the choices might be interpreted as revealed preference over the outcomes (0,0,40) vs. (80,80,0), for example, for strong inequality averse individuals, lie aversion still seems to manifest itself for a significant minority of participants.

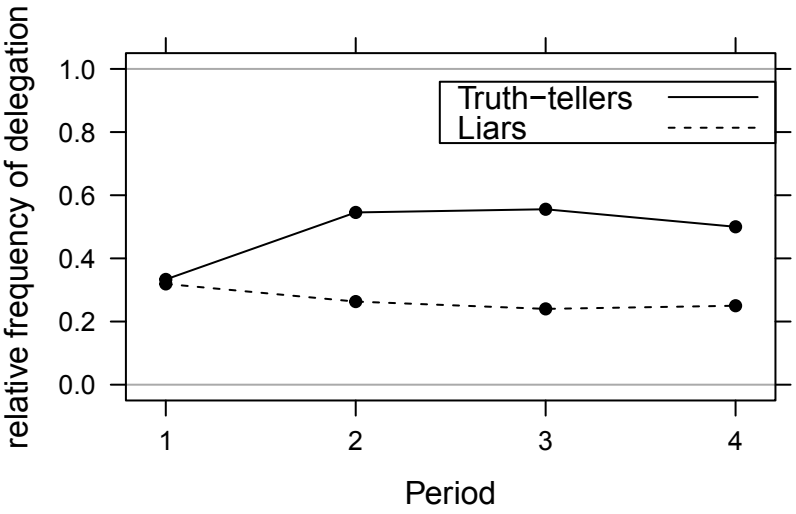


Figure 3.5: Delegation among truth tellers and liars in the "conflict" condition

egating truth-tellers and liars. In the "conflict" condition 47% of all senders who otherwise tell the truth delegate. In contrast, only 27% or all senders who otherwise lie delegate. This difference in behaviour emerges only during the experiment. The difference in delegation in the first period is not significant. To test for differences between senders who tell the truth and senders who lie also in the later periods we use a mixed effects logistic model where we include a random effect for the participant and a random effect for the matching group.

$$P(\text{delegation}) = \mathcal{L}\left(\beta_0 + \beta_1 d_{\text{Truth}} + \gamma_t + \epsilon_i + \epsilon_g\right) \tag{3.1}$$

Here $\mathcal{L}$ is the logistic function, d_{Truth} is a dummy which is one for truth tellers, γ_t is a fixed effect for period t, ϵ_i is a random effect for the individual, and ϵ_g is a random effect for the matching group.

β_1 is significantly different from zero with a p-value of 0.027. We can, thus, confirm Hypothesis 3.2.

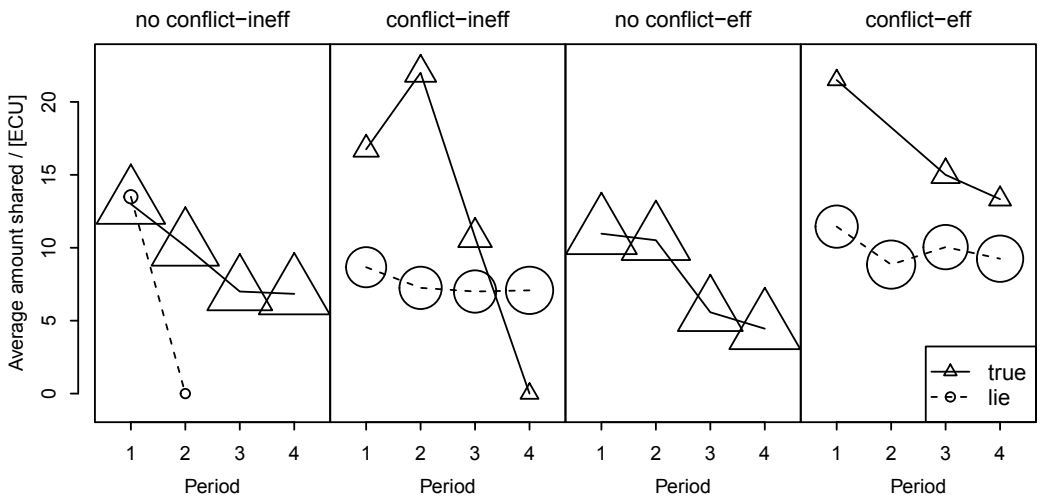
3.4.3 Compensatory behavior

We measure compensatory behavior as willingness to share money with a stranger in a dictator game. In the following we do not consider 106 decisions of senders who obtained a profit of zero in the sender-receiver game. These senders had nothing which they can share in the dictator game. We still elicited hypothetical decision for these senders to keep them busy. However, senders knew their earnings from the first round. We restrict our sample to 438 decisions of senders with positive earnings from the sender-receiver game.

First, we have found a relatively high willingness to share money: as many as 53.2% of the senders who earned 80 ECU in part 1 gave away positive amounts to the receiver. Among those senders who shared positive amounts, the average was 17 ECU or about 21.2% of the senders' earnings.

Do liars share more than truth-tellers? The average amounts shared by the senders in the two conditions are compared in Figure 3.6. The two panels in the left part of Figure 3.6 show the ineffective senders, i.e. those which were not selected for a delegation decision by the computer. The two panels on the right show the effective senders. In each group of two panels the left one shows the "no conflict" situation, the right one shows the "conflict" situation.

Interesting are, in particular, the effective players, i.e. those whose delegation decision was actually implemented. Here in the "no conflict" case (third panel from the left), all players told the truth. Shared levels are lower than in the "conflict" case (fourth panel). In the "conflict" case shared amounts are particularly high for truth telling players and intermediate (but higher than in the



Sizes (areas) of the symbols are proportional to the number of observations. Graphs only show senders with positive earnings from the first round.

Figure 3.6: Average amount shared

"no conflict" case) for liars.

Clearly, telling the truth in a situation with or without conflict are two different matters. In our experiment each player experiences both situations, players either start with conflict and conclude the experiment with no conflict or they do just the opposite. To better understand the impact of the treatments we estimate the following mixed effects regressions:

$$\text{Share} = \beta_0 + \beta_1 d_{\text{Truth}} + \gamma_t + \epsilon_i + \epsilon_g + \epsilon_{igt} \quad (3.2)$$

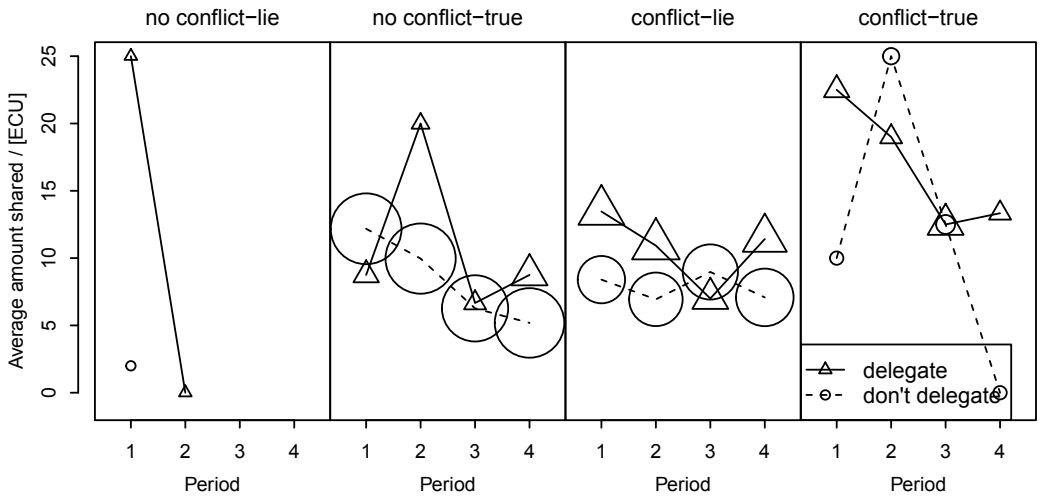
$$\text{Share} = \beta_0 + \beta_1 d_{\text{Conflict}} + \gamma_t + \epsilon_i + \epsilon_g + \epsilon_{igt} \quad (3.3)$$

d_{Truth} is a dummy which is one for senders who tell the truth, d_{Conflict} is a dummy which is one in the "conflict" condition, γ_t is a fixed effect for period t, ϵ_i is a random effect for the individual, ϵ_g is a random effect for the matching group, and ϵ_{igt} is the residual. Table 3.1 provides the estimation results. The first two columns

Table 3.1: Mixed-effect regression estimates for Equations 3.2 and 3.3

	all, Eq. (3.2)	effective, Eq. (3.2)	all, Eq. (3.3)	effective, Eq. (3.3)
Hypotheses				
Truth	−1.78**	−2.21*		
	(0.66)	(1.07)		
Conflict			1.80**	2.90**
			(0.62)	(1.04)
Control				
Period 1	12.61***	13.21***	10.75***	10.64***
	(1.07)	(1.45)	(1.01)	(1.34)
Period 2	10.59***	10.11***	8.73***	7.64***
	(1.04)	(1.42)	(1.01)	(1.37)
Period 3	9.25***	10.11***	7.25***	7.26***
	(1.06)	(1.46)	(1.02)	(1.43)
Period 4	7.82***	7.85***	5.95***	5.23***
	(1.03)	(1.40)	(0.99)	(1.32)
Num. obs.	438	219	438	219
Num. groups: i	136	118	136	118
Num. groups: g	7	7	7	7

Profile Likelihood 0.95 confidence intervals in brackets. * 0 outside the confidence interval.



Sizes (areas) of the symbols are proportional to the number of observations. Graphs only show senders with positive earnings from the first round.

Figure 3.7: Average amounts shared by delegation choice.

in Table 3.1 provide estimation results for Equation (3.2). We find that senders who tell the truth share an amount significantly smaller than those who lie — regardless whether we consider all senders with a positive income from the first round or only the effective senders. This supports Hypothesis 3.3.

The two columns on the right of Table 3.1 provide estimation results for Equation (3.3). We find that in both cases, all senders and effective senders, the coefficient for the "conflict" condition is significantly different from zero. This supports Hypothesis 3.4.

Does delegation produce less compensation? Figure 3.7 compares the sent amounts for senders who delegate with those who do not delegate in different situations. Most interesting is the third panel: participants in the "conflict" treatment who "lie". In contrast to our Hypothesis 3.5 senders who delegated share more than those

Table 3.2: Mixed-effect regression estimates for Equation 3.4

	all conflict, Eq. (3.4)	effective conflict, Eq. (3.4)
Hypothesis		
Delegation	4.65 (1.78)**	5.47 (2.44)*
Control		
Period 1	9.96 (1.76)***	10.69 (2.45)***
Period 2	8.39 (1.66)***	7.46 (2.38)**
Period 3	8.76 (1.60)***	9.54 (2.24)***
Period 4	6.65 (1.60)***	7.27 (2.22)**
Num. obs.	204	102
Num. groups: i	128	80
Num. groups: g	7	7

Profile Likelihood 0.95 confidence intervals in brackets. * 0 outside the confidence interval.

senders who have chosen to send the message themselves.

To more precisely assess the effect of delegation we estimate the following mixed effects regression:

$$\text{Share} = \beta_0 + \beta_1 d_{\text{Delegation}} + \gamma_t + \epsilon_i + \epsilon_g + \epsilon_{igt} \tag{3.4}$$

$d_{\text{Delegation}}$ is a dummy which is one for senders who delegate, γ_t is a fixed effect for period t, ϵ_i is a random effect for the individual, ϵ_g is a random effect for the matching group, and ϵ_{igt} is the residual. Estimation results are shown in Table 3.2. Regardless whether we look at all senders in the "conflict" treatment or only at the effective senders: senders who delegate share, on average, substantially higher amounts than those who do not delegate. It is hence not, as we had hypothesized above in Hypothesis 3.5, that senders who delegate share less. Instead, the delegation opportunity seems rather to serve as a screening device for senders with different psychological costs of lying. Senders with a low cost of lying don't have to compensate their lies. They neither delegate nor compensate through sharing in the subsequent dictator game.

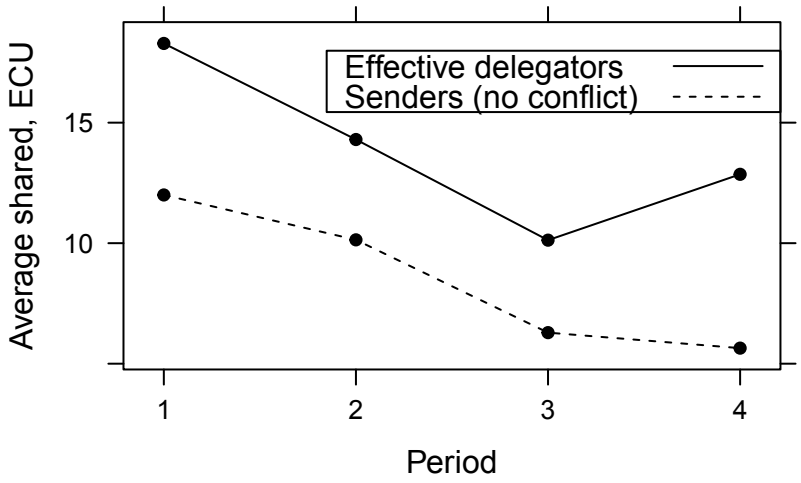


Figure 3.8: Sharing: Effective delegators in the "conflict" vs. senders in the "no conflict" condition

Senders with a high cost have to use both instruments to reduce this cost: they have to delegate and, in addition, they also have to compensate in the subsequent dictator game.

Importantly, participants who effectively delegated sending the message in the "conflict" condition (have chosen "delegate" and were selected by the computer) also shared more than participants in the "no conflict" condition (See Figure 3.8). To confirm this we again estimate Equation 3.4 and include only effective senders with positive earnings from the first round. As in Figure 3.8 we restrict to either 'no conflict' or delegators from 'conflict'. Results are shown in the left column of Table 3.3. We find that the coefficient for "Delegation" is significant and positive.

To more closely assess the effect of responsibility we estimate Equation (3.4) again for effective senders with positive earnings from the first round. Now we include only senders from the "con-

Table 3.3: Mixed-effect regression estimates for Equation 3.4

	conflict deleg./no conflict	lie delegation
Hypothesis		
Delegation	4.69 (1.40)***	3.23 (2.33)
Control		
Period 1	12.53 (1.27)***	8.28 (2.36)***
Period 2	10.33 (1.24)***	8.09 (2.15)***
Period 3	7.64 (1.25)***	5.89 (2.08)**
Period 4	3.96 (1.26)**	5.29 (2.05)**
Num. obs.	304	91
Num. groups: i	136	73
Num. groups: g	7	7

Profile Likelihood 0.95 confidence intervals in brackets. * 0 outside the confidence interval. Both models include only effective senders with positive earnings from the first round. The model "conflict deleg./no conflict" includes all 'no conflict' and only delegators from 'conflict', the model "lie delegation" includes only "conflict" and only if the outcome was 'lie'.

flict" condition where the outcome was "lie", either through a direct lie or through an indirect lie, i.e. through delegation. Results are shown in the right column of Table 3.3. We find that the coefficient for "Delegation" is positive, though not significantly different from zero.

3.5 Conclusions and Discussion

We allowed senders to delegate the decision to others with perfectly aligned interests and informed receivers about this opportunity. Although we create strong incentives to lie directly, a significant share of senders decides to delegate. Among various explanations of why senders delegate, 'distancing from the harm' remains a promising candidate.

Our results add to the discussion of lying aversion and suggest

that for some people it is not the losses to the affected party (in our setting net social effect of lying was positive) but rather the directness of lying senders are trying to avoid.

In our setting, many senders who prefer to delegate would chose a truthful message otherwise. This suggests that institutions which allow for delegation and where delegation looks innocent, could actually do better if delegation was restricted.

In line with the literature on moral balancing, we find that lying generates compensatory behavior. After a lie participants give away a larger fraction of their earnings to another player with whom they did not interact before. Participants who told the truth share less. At the same time we also observe consistency: participants who are ready to sacrifice money through truth-telling in the sender-receiver game are also more ready to share in the dictator game.

Perhaps most interestingly, indirect liars (delegators) share more (and not less) than direct liars if there is a conflict between senders and receivers. As a explanation we suggest that delegation serves as a screening device: people with a low psychological costs of lying select into direct lying. People with a high psychological cost of lying reduce this cost at least partially by delegation. To reduce this cost furthermore they also share more in the dictator game.

3.6 Appendix: Experimental Instructions

Welcome and thank you for participating in the experiment of Max Planck Institute for Economics. Please switch off your mobile phones and remain silent. All the decisions you make in this experiment will be treated anonymously. Any communication between the participants is not allowed. Should you have any questions during the experiment, please raise your hand, and one of the experimenters will answer you privately. No participant will learn the indentity of any other participant.

Your earnings in this experiment depend on the decisions you make, the decisions of the other participants as well as on chance. During the experiment your earnings will be expressed in ECU (Experimental Currency Units) which will be translated into EUR at the end of the experiment at the following exchange rate:

10 ECU = 1 EUR.

In addition to your earnings in the experiment you will receive a fixed show-up fee of 2.5 EUR. Your total earnings (including show-up fee) will be paid out to privately at the end of the experiment.

Remarks for the experiment

There are three roles in this experiment: Player 1, Player 2 and Player 3. At the beginning of the experiment one third of the participants will be randomly assigned to the role of Player 1, one third to the role of Player 2, and one third to the role of Player 3. Each participant will keep her role throughout the experiment.

The experiment lasts for four periods. Each period consist of two parts. The following instructions describe the procedure in part 1; the instructions for part 2 will be distributed once you finish part 1 in period 1 of the experiment. The procedure in Part 1 and Part 2 in each of the periods (unless specified otherwise) is identical.

At the end of the experiment one of the periods will be randomly chosen for payment. Your payment depends on both parts. Since each period has an equal chance to be selected, please decide carefully.

At the end of the experiment you will be asked to answer a questionnaire. Your answers do not affect your earnings.

Instructions for Part 1

At the beginning of part 1 computer forms groups of three participants: each group consists of one Player 1, one Player 2 and one Player 3. The groups will be then assigned with equal probability to one of the two experimental conditions: **Condition 1** or **Condition 2**. In the first period 50% of the groups are assigned to Condition 1 and 50% of the groups are assigned to Condition 2 of the experiment.

In the second period, **Player 1** and **Player 2** interact in the same condition they were assigned to in the first period of the experiment. In the remaining two periods of the experiment (Period 3 and Period 4), Player 1 and Player 2 change the condition so that each Player 1 and Player 2 participate twice in each of the two conditions.

Player 3 is assigned to either Condition 1 or Condition 2 with the same probability in each period of the experiment.

Before Player 1 and Player 2 make their decisions they will be informed about which condition they are assigned to; Player 3 will be informed about the condition he is assigned to only after he makes his decision.

Participants in both conditions will be asked to make their choices in the following task. In the task there are five virtual boxes (numbered 1 to 5) that might contain monetary prizes. The value of the prize is fixed: **80 ECU** for Player 1, **80 ECU** for Player 2, and **40 ECU** for Player 3. The allocation of prizes into boxes depends, however, on the condition participants are assigned to.

In **Condition 1** computer randomly selects one of the five boxes and puts the prizes for all the three players into that Box. Please note that for all groups in this condition it is always true that the prizes for all the three players will be in the same Box. The other four boxes will not contain any prizes.

In **Condition 2** computer randomly selects two of the five boxes and puts the prizes for Player 1 and Player 2 into one box, and the prize for player 3 into the other box. Please note that for all groups in this condition it is always true that the prizes for Player 1 and Player 2 will be in the same box, but the prize for Player 3 will be in a different Box. The other three boxes will not contain any prize.

The earnings for the three players in the group depends on the choice of Player 3. Player 3 will be asked to open a box (without knowing the allocation of the prizes). Before making his decision Player 3 will receive a message either from Player 1 or from Player 2: 'Your prize is in Box x'.

In contrast to Player 3, Player 1 and Player 2 will see on their screens the actual allocation of prizes. They will be therefore in-

formed about the condition they are assigned to. Player 1 and Player 2 will then be independently asked to specify the number of the box they advise Player 3 to open: 'Your prize is in Box x'. Together with the number of Box they would like to advise, both Player 1 and Player 2 are asked to indicate if they would like to send *my own message* or *'the message by Player 2(1)'* respectively. After Player 1 and Player 2 have made their decisions, computer selects either Player 1 or Player 2 with equal probability and implements his decision: if the selected player has chosen the other player to send a message, the number of the box advised by the other player is sent to Player 3; if the selected player has chosen *'my own message'*, the number of the box advised by that player is sent to Player 3.

Player 3 then receives a message that says 'Your prize is in Box x' with the information about who is sending the message. Player 3 will thus read:
'Player (1) 2 has chosen to send the message herself. Player (1) 2 sends you the following message' (in case the selected player has chosen *'my own message'*
or

'Player 1(2) has chosen Player 2(1) to send you the message' (in case the selected player has chosen *'The message by Player 2(1)'*.

Player 3 will be asked to open a box and define the earnings for the players in his group.

On the next page you will see the examples of the information that players receive before making their decisions.

Example* for the information that Player 1 and Player 2 receive in **Condition 1**:

Box 1	Box 2	Box 3	Box 4	Box 5
Prize1: 0 Prize2: 0 Prize3: 0	Prize1: 80 Prize2: 80 Prize3: 40	Prize1: 0 Prize2: 0 Prize3: 0	Prize1: 0 Prize2: 0 Prize3: 0	Prize1: 0 Prize2: 0 Prize3: 0

Example** for the information that Player 1 and Player 2 receive in **Condition 2**:

Box 1	Box 2	Box 3	Box 4	Box 5
Prize1: 0 Prize2: 0 Prize3: 0	Prize1: 0 Prize2: 0 Prize3: 40	Prize1: 0 Prize2: 0 Prize3: 0	Prize1: 80 Prize2: 80 Prize3: 0	Prize1: 0 Prize2: 0 Prize3: 0

Information that Player 3 receives in **Condition 1** and in **Condition 2**:

Box 1	Box 2	Box 3	Box 4	Box 5
Prize1: ? Prize2: ? Prize3: ?	Prize1: ? Prize2: ? Prize3: ?	Prize1: ? Prize2: ? Prize3: ?	Prize1: ? Prize2: ? Prize3: ?	Prize1: ? Prize2: ? Prize3: ?

*,** These are only examples for many possible prize allocations. The exact allocation of prizes in the experiment will be defined as described above.

After Player 3 has made his decision, all players will be informed about the following: the allocation of prizes, the message and the sender of the message Player 3 received, the decision of Player 3, and the earnings of the participants in their group.

The experiment will then continue to Part 2.

If you have read the instructions carefully and do not have further questions, please click on 'Continue'. Before the experiment starts, you will be asked a few control questions to make sure you understand the course of the experiment.

Good luck!

Additional Instructions for Part 2

At the beginning of Part 2 of the experiment the computer forms new groups of three: Player 1, Player 2, and Player 3. Each participant keeps his role from part 1 of the experiment. Matching procedure guarantees that no player is matched with the same participant he interacted with in the first part of the experiment.

After the groups have been formed Player 1 and Player 2 will be independently asked how much of their earnings from part 1 of the experiments they would like to transfer to Player 3. This Player 3 is *NOT* the Player 3 with whom they interacted in Part 1 but a new participant that was a Player 3 in another group. Player 1 and 2 can transfer any amount between 0 and 80 ECU (in case the players receive 0 ECU in Part 1 they will be asked how much they

would transfer should they have earned 80 ECU). When Player 1 and Player 2 make their decision they know only their own earnings from Part 1. No one knows the earnings of the other players.

Once Player 1 and Player 2 have made their decisions, the computer randomly selects one of the two players. Only the decision of the selected player will be implemented and the respective amount will be transferred from that player to Player 3. The decision of not selected player (and the decision of players with 0 ECU from Part 1) will not be implemented. The probability that Player 1 or Player 2 is selected is equal.

The participants will then be informed about the following:

- the amount that Player 1 and Player 2 have chosen to transfer;
- the randomly selected player;
- the total earnings of all the players in the group.

Part 2 of the experiment is then over.

If you have read the instructions carefully and have no further questions, please click on 'Continue' to start Part 2 of the experiment.

Chapter 4

Moral Substitutes and Moral Complements: the Effects of Third-Party Action on Moral Behavior

4.1 Introduction

4.1.1 Moral Consistency and Moral Balancing

The path dependency of ethical behavior is a robust phenomenon in the psychological literature. The positive correlation between past and future ethical actions, coined as *moral consistency*, has been a dominating paradigm in understanding moral behavior in a dynamic context. People are believed to exhibit consistency to reduce the cognitive dissonance (Festinger, 1962), to maintain their self-image (moral identity models, Akerlof & Kranton (2000)) or to comply with social norms (Andreoni & Bernheim, 2009). "Foot-in-the door" social compliance technique (Freedman & Fraser, 1966) is a prominent example of utilizing people's longing for consistency in

a prosocial domain: small donation requests increase the commitment to larger donation requests in the future (Schwarzwald *et al.*, 1983).

The positive correlation between past ethical actions and the subsequently shown level of altruism has found support in recent experimental studies. In Gneezy *et al.* (2012a) a donation task is followed by a sender-receiver game. Consistent with the idea of a "foot-in-the door" strategy to foster prosocial behavior, they observe that having just made a small donation increases the level of truth telling. Gino *et al.* (2010) prime participants as being of an ethical or unethical type by merely assigning them original or counterfeit products. They then study their subsequent morally relevant behavior. The authors observe that the participants primed for being of an unethical type were more dishonest compared to the ethically primed group. In a recent study, Cojoc & Stoian (2014) combine a lying task with the possibility of a subsequent donation to charity. By manipulating the available information about the possibility to donate, they find support for the "conscience numbing" hypothesis: the first unethical action inhibits the marginal disutility of the subsequent unethical action which ultimately leads to consistent moral behavior.

As opposed to moral consistency, people may also exhibit a negative correlation between past and future moral actions, usually described as moral inconsistency or *moral balancing*, as recent theories and experiments suggest (Nisan & Horenczyk, 1990; Meritt *et al.*, 2010; Mazar & Zhong, 2010). According to the moral balancing hypothesis, having done good in the past (earning moral credits) can license transgressions, whereas doing bad (earning moral debits) calls for subsequent cleansing. Both sides of the proposed compensatory behavior, "moral licensing" and "moral cleansing", have

found empirical support.

In Mazar & Zhong (2010) participants who bought green products and presumably felt morally superior shared less money in a dictator game and cheated more towards the experimenter compared to buyers in a conventional online store. After imagining volunteering for a public project, participants were more likely to choose luxury instead of necessity products (Kahn & Dhar, 2006) or cheated more (Clot *et al.*, 2014). Participants primed with positive words ("helpful", "generous") donated less compared to a neutral and negative traits condition (Sachdeva *et al.*, 2009). This could be interpreted as the dark side of moral balancing, since ethical deeds in the past license moral transgressions in the future.

Participants, on the other hand, who did not return an extra payment to the experimenter and presumably went low on their moral account were more likely to make a real donation to a charity (Gneezy *et al.*, 2012b). Participants who recalled recent unethical acts were striving to compensate for it by a higher propensity to donate or volunteer (Jordan *et al.*, 2011). This behavior could be seen as the light side of moral balancing, since unethical deeds in the past are compensated by praiseworthy behavior in the future.

Given evidence for both behavioral patterns, researchers discuss under which conditions consistent behavior and under which conditions balancing behavior emerges. Gneezy *et al.* (2012a) compare costless and costly donations and argue that higher costs provide a better signal to one's true preferences, thus producing consistency. Conway & Peetz (2012) and Jordan *et al.* (2011) suggest that the higher level conceptual abstraction (distant vs. recent actions) prompts consistency.

Cornellissen *et al.* (2013) propose an interesting synthesis of both concepts. They argue that the mind set, being either outcome-

based or rule-based, explains the direction of the link between past and future moral actions. In their study, participants are categorized according to their mind set on the basis of a standard trolley dilemma question. Outcome-based participants tend to balance their recalled past unethical behavior with a higher subsequent donation in a dictator game and their past ethical behavior with a lower donation. This is in sharp contrast to rule-based participants who strive for consistency: they increase donations after having recalled an ethical act, and decrease them after recalling an unethical one.

4.1.2 The Effect of Third-Party Actions

As the literature review in the previous section demonstrates, there is an intense discussion about possible effects of the path dependency of one's own moral actions. The impact of observed moral actions of a third party on one's own subsequent moral actions, however, has received less attention. This question is, however, important since a person rarely acts as Robinson before the arrival of Friday. She is rather embedded in a nexus of social interactions.

The influence of social environment on the individual behavior has a long tradition in sociology. A "broken windows" paradigm suggests that small damage by others (a broken window, a graffiti on the wall) can signal social tolerance towards even further destruction (Wilson & Kelling, 1982).

Experimental literature confirms that (the information about) the behavior of others with no stake in the game has a substantial affect on the behavior of the actors.

Engel *et al.* (2011) provide pre-play examples of group behavior in public good games. They observe that informing partici-

pants about highly cooperative groups in previous experiment enhances contributions. Similarly, giving participants information about poorly cooperative groups dramatically lowers contributions.

Rauhut (2013) investigates the effect of positive and negative examples on incidence of lying. He observes that providing information about the rate of lying by others affects the degree of lying of actors. Participants who learn that the others lie less increase their truth-telling, and participants who learn that others lie more seem to overreport more. He suggests that the changes in behavior are due to differences in prior (and subsequent adjustment) of beliefs about the social norm.

Providing positive examples is also considered as possible instrument in tax compliance policy toolkit. However, the evidence about the direction and the strength of these effects is mixed. In some experiments, observing high compliance rates increases one's own compliance. In other studies, only negative examples (low tax compliance) have an impact on one's own compliance (See Alm *et al.* (2013) for a discussion).

In the moral domain, the presence of third-party actions raises an interesting question. Benevolent actions by others may lead to a free-riding problem, where the good actions of others, i.e., providing help to those in need, substitute for one's own willingness to behave prosocially Buchanan (1975). This has been referred to as the Samaritan's dilemma. It can easily be explained by considering ethical actions as a public good which is systematically underprovided due to the characteristic free rider problem.

An alternative approach would be to consider third-party actions as non-substitutes or even as complements. Andreoni's idea of a "warm glow" that subjects experience when engaging in prosocial behavior considers ethical deeds rather as a matter of private

consumption (Andreoni, 1990). Ethical third-party actions may then rather serve as an inspiration and accordingly a complement for own ethical behavior than as a substitute.

We believe that Cornellissen *et al.*'s elicitation of people's mind set is helpful in understanding the link between one's own and others' moral behavior. Outcome-minded participants may feel that their own behavior somehow adds to the aggregate consequences and therefore adjust it in order to compensate for others' "under-" or "overcontribution" to a moral aim. Rule-minded participants who are oriented toward following a moral standard will rather take the observed behavior of others as a norm to follow, thereby aggravating the trend of an "under-" or "overcontribution".

4.2 Research Question and Hypotheses

In a baseline, we reproduce the experiment of Cornellissen *et al.* (2013) by priming subjects with an ethical or unethical recall stage, respectively, in order to investigate whether they exhibit moral balancing or moral consistency. In particular, in line with their study, we classify subjects depending on their mind set according to their answer to a standard trolley problem into outcome-oriented and rule-oriented participants.

Side Hypothesis: In the baseline, we expect to replicate Cornellissen et al.'s results by observing moral balancing for outcome-based participants and moral consistency for rule-based participants.

We then analyze, whether third-party actions serve as substitutes or complements of own ethical behavior. In this respect, we expect that it is important that the third-party action has moral relevance.

This is likely to be the case, if it is caused by another human subject. To control for this, we check whether actions imposed by random assignment are already sufficient to cause a difference. We expect that this is not the case.

Control Hypothesis: Third-party actions imposed by random assignment have no effect on subjects' own moral actions

We conjecture that outcome-based participants are likely to view the actions of others as substitutes for their own actions, since they are mainly focused on the unfolding effects of ethical actions. In particular, they gain no benefit from executing the action per se while following a particular behavioral rule. This is different for rule-oriented participants which are likely to consider the actions of others as complements. Since their focus is on moral standards incorporated in rules of action, subjects living up to these rules will be likely to reinforce their own rule-oriented behavior. They will then be inspired by an ethical third-party action and follow its example. If that is true, the influence of a third-party action on outcome-oriented and rule-oriented participants' moral actions is systematically different between outcome-based and rule-based participants. To test for this proposition, we influence the behavior of ethically primed participants with an ethical third-party action and the behavior of unethically primed participants with an unethical third-party action. Subsequently, we compare the ethical behavior after this manipulation with the one observed in the baseline. Our hypotheses on the impact of ethical third-party behavior are therefore as follows.

Main Hypothesis 1: Outcome-oriented participants will consider

moral third-party actions as substitutes for their own actions.

Main Hypothesis 2: Rule oriented-participants will consider moral third-party actions as complements for their own actions.

Ethically primed outcome-based participants who are confronted with an ethical third-party action will tend to give less, because the third party is already acting morally on their behalf. Unethically primed outcome-based participants confronted with an unethical third-party action, on the other hand, will tend to give more to compensate for this third-party behavior. These contravening tendencies will consequently mitigate the moral balancing effect observed for outcome-based participants in the baseline.

Conversely, ethically primed rule-based participants confronted with an ethical third-party action will tend to give more, because the third party leads by good example. Unethically primed rule-based participants who are confronted with an unethical third-party action will analogously tend to give less, because the third-party serves as a demoralizing example.

4.3 Experiment Design

We employ a 2 x 2 factorial design. The variables that we manipulated in the conditions of the experiment are current moral account (manipulated by recalling ethical or unethical act) and the source of good/bad action towards the participants (other subject or experimenter). Table 1 gives and overview over the experimental conditions and the resulting treatments.

The experiment consists of three parts: recall task, dictator game, and real-effort task. Participants are informed about the

	Action: Random assignment	Action: Third-party
Priming: Ethical	low or high effort	low effort
Priming: Unethical	high or low effort	high effort

Table 4.1: Experimental Treatments

three parts of the experiment, but receive the instructions for the dictator game only after they have finished the recall task (see the instructions).

At the beginning of the experiment, half of the participants are assigned to the "Ethical" and the other half to the "Unethical" Condition. Depending on the condition, subjects are consequently asked to recall a recent ethical or unethical act, describe the action itself and the consequences of that action for the affected person[1].

The middle part of the experiment consists of a dictator game. In the baseline, all subjects are paired and play this dictator game. In the third-party treatment, two subjects are randomly chosen in the beginning of this part to be in the role of effort allocators, one in each condition. These two subjects do not play a dictator game. Dictators are endowed with 10.00€ and are asked how much they want to transfer to a recipient. We elicit dictator choices from both the subjects in the pair and then randomly select one of the subjects to be a dictator and implement his decision.

In the real-effort task, participants face a 10 x 10 table containing zeros and ones in randomized order and have to report the frequency of ones in that table (with a tolerance of ± 1). Participants learn about the content of the real effort task and are told that they will

[1]These features of the recall task have been shown to prime compensatory behavior (Conway & Peetz, 2012; Cornellissen *et al.*, 2013).

be informed about how many tables they will have to solve after the recall task: either **one** ("low effort") or **eight** tables ("high effort").

In the baseline, participants know that the number of tables to solve is randomly assigned to them. In the third-party treatment, participants know that the number of tables to solve will be allocated by another participant. For this manipulation, we randomly selected the two effort allocators in a session to decide about the number of tables. In the "Ethical" Condition an effort allocator is strongly incentivized to choose a default option of **one** table. In the "Unethical" Condition an effort allocator is strongly incentivized to choose a default option of **eight** tables. This means that an effort allocator would loose his whole endowment of 10.00€, if he deviates from the default.

The other participants do not learn about the incentives of the effort allocators, but are simply informed how many tables their group members and how many tables the other group members receive from an effort allocator (see instructions).

After receiving the information about the number of tables to solve in the last part of the experiment, participants play the dictator game. Therefore, they may be influenced by this information in their choice of which fraction of their endowment to transfer to the other participant. Remember, that in the dictator game, both participants of each pair decide for the role of dictators. The computer then chooses one of the two decisions to be implemented with equal probabilities.

Figure 4.1 summarizes the flow of the experiment.

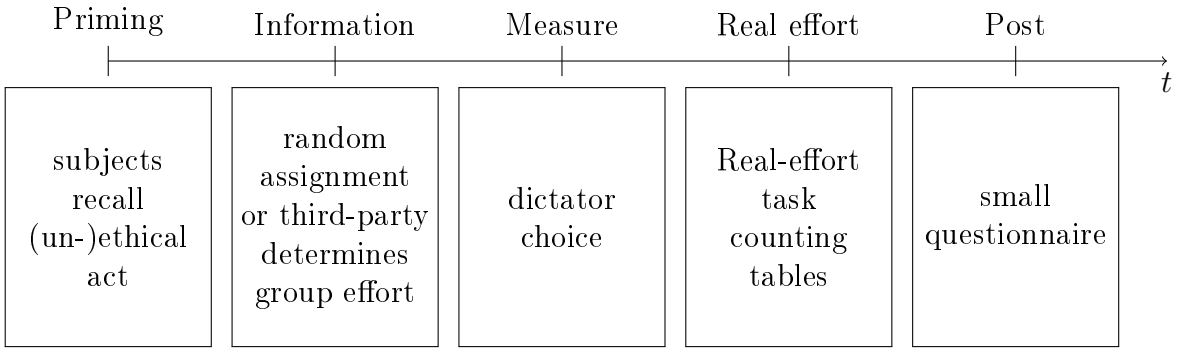


Figure 4.1: Flow of the experiment

4.4 Results

We ran seven experimental sessions in August 2014 in the Experimental Laboratory of the Max Planck Institute of Economics in Jena, Germany. The experiment was programmed in z-Tree (Fischbacher, 2007), the participants were recruited with ORSEE (Greiner, 2004). A session lasted approximately 50 minutes, and the average payment was 80 ECU (10 ECU=1 EUR) (including show-up fee).

We split our sample into outcome-based and rule-based participants. Similarly to Cornellissen *et al.* (2013), the participants are categorized by their response to a standard trolley dilemma. Those who said that it was acceptable to pull the lever to save five people and kill one person are referred to as "outcome-based" in our analysis. Those who said that it was unacceptable are referred to as "rule-based'.

4.4.1 Recalling Ethical or Unethical Actions

Overall, we observed a relatively high willingness to share money. 164 out of 218 participants (more than three quarters) shared positive amounts in the dictator game, while 54 participants did not share anything. The aggregated average amount shared was 25.4 % of the participants endowment of 10.00 EUR. These findings are consistent with the overall generosity found in other dictator game experiments (see Engel (2011) for a meta-study comparison).

We start by turning our attention toward our first hypothesis and analyze if recalling a recent ethical or unethical action (i.e., an action with positive or negative consequences for another person, respectively) had any effect on subsequent giving in a dictator game.

First, we compare the average amount shared and the fraction of selfish dictators (a subject giving zero in the dictator game) in the two conditions. The average amount shared in the ethical condition is 5.00 ECU higher compared to the unethical condition (29.58 ECU vs. 24.55 ECU, $p = 0.117$ according to a two-sided Wilcoxon rank sum test).

The fraction of selfish dictators is twice as large in the unethical condition. In the ethical condition, 7 out of 64 (11 %) chose to give zero to the recipient, in the unethical condition 16 out of 64 participants (25 %) did so ($p = 0.064$ according to the Fisher exact test). Given these observations, we conclude that the recall of ethical or unethical past actions affected the willingness to donate itself rather than the amount donated.

Figure 4.2 illustrates subjects' reactions to the ethical and unethical recall priming. We observe that outcome-based participants react moderately to this priming: ethically primed subjects give an average amount of 29.66 ECU, while the unethically primed ones

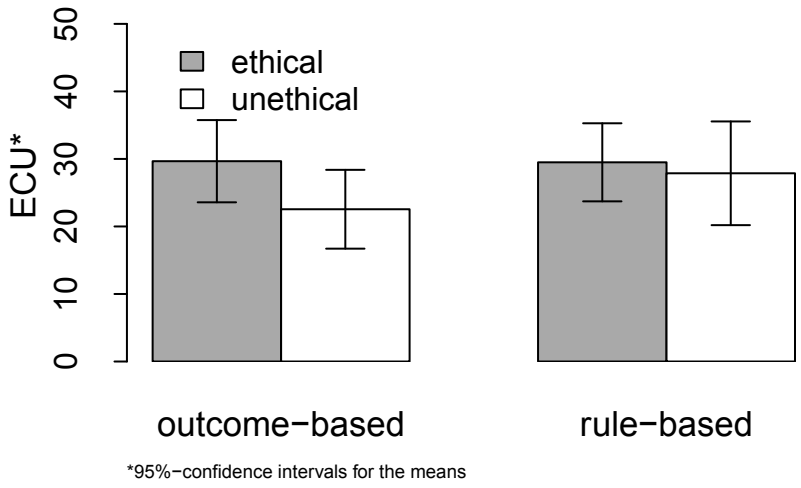


Figure 4.2: Average amount shared in the Baseline by condition and mind-set, ECU

give 22.55 ECU ($p = 0.093$). Rule-based participants, on the other hand, show no reaction at all: ethically primed subjects share on average 29.50 ECU, while the unethically primed ones share 27.87 ECU ($p = 0.767$).

As far as outcome-oriented subjects are concerned, these findings are in sharp contrast to Cornellissen *et al.* (2013), since recalling a past ethical act increases and not decreases subsequent giving. Therefore, we do not find evidence for moral balancing but rather for moral consistency in subjects with an outcome-based mindset. In this sense, not only suffering monetary costs (as argued by Gneezy *et al.* (2012a)) but also remembering non-monetary costs might trigger generosity in people with an outcome-based mindset. Since we expected to replicate the results of Cornellissen, our side hypothesis is falsified.

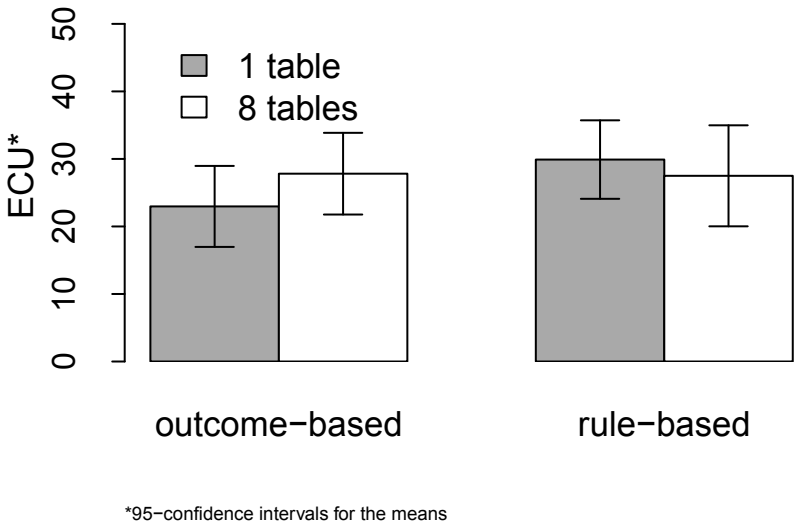


Figure 4.3: Sharing in the Baseline by the number of tables, ECU

4.4.2 Low and High Efforts by Random Assignment

We assume that the mere amount of tables that subjects have to solve does not affect their behavior. In particular, if efforts are simply randomly assigned to the participants, we presume that this has no significant impact on participants' behavior. In this sense, it is crucial that the third party is taking a deliberate action. Figure 4.3 presents the average amounts shared by outcome-based and rule-based participants in the baseline where moral connotations are excluded via a random assignment of effort.

As Figure 4.3 shows, the mere number of tables the participants have to solve does not affect the amount shared in the dictator game ($p = 0.285$ for outcome-based and $p = 0.757$ for rule-based participants according to a Wilcoxon rank sum test). This means that observing "fate" or coincidence does not have an increasing or decreasing effect on moral behavior. Our control hypothesis is

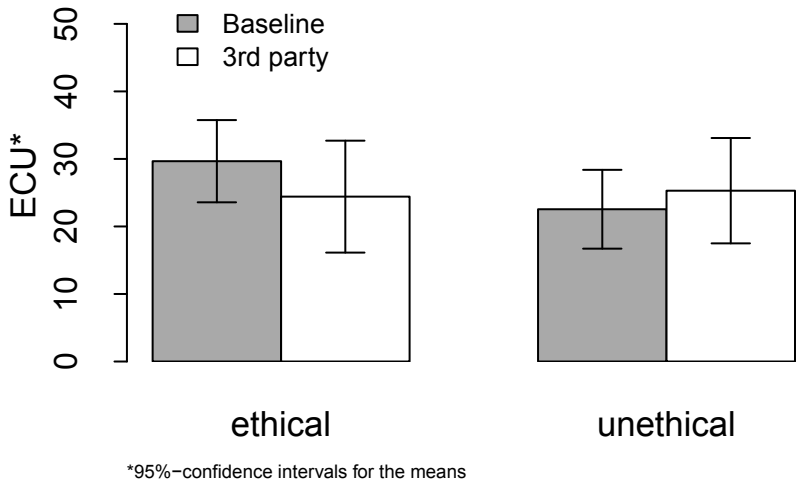


Figure 4.4: Sharing by outcome-based participants, ECU

therefore corroborated.

This result allows us to pool the data with low and high efforts in the baseline for the further analysis.

4.4.3 Low and High Efforts by Third-Party Choice

We now turn to analyze the differences in sharing between the baseline and the third-party treatment. Figures 4.4 and 4.5 compare the observed sharing by outcome- and rule-based participants in the baseline and in the manipulation in which a third party intervened.

As Figure 4.4 shows, ethically primed participants which are outcome-minded are not affected by observing an ethical third-party action. They share 29.67 ECU in the baseline and 24.42 ECU in the third-party treatment ($p=0.571$, Wilcoxon rank sum test). Neither are unethically primed participants which are outcome-minded influenced by observing an unethical third-party action.

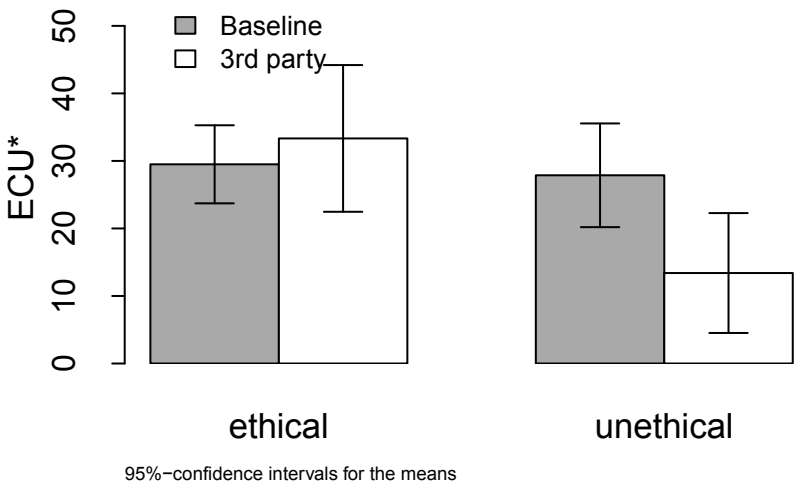


Figure 4.5: Sharing by rule-based participants, ECU

They share 22.55 ECU in the baseline vs. 25.30 ECU in the third-party treatment (p=0.377, Wilcoxon rank sum test).

Our first main hypothesis is falsified. Outcome-minded subjects do not consider third-party actions as substitutes for their own actions. In fact, third-party actions do not affect outcome-minded participants behavior in any direction. Thus, outcome-minded participants react symmetrically insensitive to moral third-party actions.

As Figure 4.5 shows, ethically primed participants which are rule-minded are also not affected by observing an ethical third-party action. They share 29.50 ECU in the baseline and 33.33 ECU in the third-party treatment (p=0.319, Wilcoxon rank sum test). Unethically primed participants which are rule-minded are, however, highly influenced by observing an unethical third-party action. They share 27.88 ECU in the baseline and only 13.40 ECU in the third-party treatment (p=0.028, Wilcoxon rank sum test).

Our second main hypothesis is partially confirmed. Unethically primed participants with rule-based mind set consider unethical third-party action as complements to their own behavior. Ethically primed participants with this mind set, on the other hand, do not adjust behavior to the positive examples set by others. In this sense, rule-minded subjects, react asymmetrically to demoralizing examples in adjusting their behavior to the negative.

4.5 Conclusion and Discussion

We explored the path dependency of moral behavior by asking participants to recall an ethical or unethical task before sharing money in a dictator game. Contrary to our primary expectations which were based on the findings by Cornellissen *et al.* (2013), we observed a consistent and not compensatory effect of the recall stage on the subsequent giving in the baseline. Participants recalling an ethical deed gave marginally more than participants in the unethical group. Our results concerning the dynamics of moral behavior therefore resonate with the results of the experiments demonstrating consistency between past ethical and unethical decisions and the subsequent prosocial behavior.

In the treatment, we interfered into the dynamics of participants' own behavior with a good or bad action by others and investigated how this action affected participants' giving in a subsequent dictator game. With this manipulation we extent the literature on the effects of the behavior of others by disjoining the domains of others' and own behavior. We investigate how the variable of interest, the amount donated, is affected by the level of efforts set by another participant. In this sense, we explore the 'generalizability' of ethical and pro-social behavior across domains addressed in studies on

negative reciprocity (Houser *et al.*, 2012).

We observed a difference in responses to others' behavior by outcome-based and rule-based participants. Outcome-based participants did neither react to ethical nor unethical actions that they observed, and were therefore symmetrically unaffected by observing moral third-party actions. Rule-based participants, on the other hand, showed a significant adjustment to unethical third-party actions by following their example. These participants treated unethical third-party actions as complements and accordingly lowered their own giving in the subsequent dictator game. This was, however, not the case for rule-minded participants observing an ethical third-party action. Rule-minded participants are thereby exhibiting an asymmetric reaction to moral third-party action by following demoralizing while ignoring commendable examples.

Our findings suggest that the distinction of participants according to the mind set as suggested by Cornellissen *et al.* (2013) is also useful for understanding how moral actions of others affect one's own subsequent behavior. In particular, we found that people may react to the observation of unethical third-party actions, if they have a rule-based mind set. They are, however, much more insensitive to ethical examples of third-party behavior, irrespective of their mind set.

The interpretation of our findings requires certain precaution. First, the insensitivity of outcome-minded participants might result from the fact that the intervention of the third-party did not impose any monetary consequences on the participants. This may be a reason that the amount of efforts to be exerted seems relevant for rule-minded participants only.

Another important limitation of the experiment is the restriction of examples to positive for ethically and negative for uneth-

ically primed participants. It may, for instance, be interesting to investigate the influence of unethical third-party actions on people memorizing ethical deeds. Is the effect of the unethical intervention sufficiently strong to inhibit consistent ethical behavior?

We believe, however, that our findings provide implications for policy measures based on informing people of socially preferable or socially condemnable behavior. In particular, the distinction of subjects into outcome-minded and rule-minded ones might be useful in the tax compliance domain where "peer effects" have been suggested to decrease tax evasion. In a recent experiment, Casal & Mittone (2014) discuss the effectiveness of social blaming (disclosed identity of tax evaders) in increasing the rate of truthful reports. Our results call for more precaution. Although psychological costs of blaming might *ex ante* prevent potential evaders from cheating, negative *ex post* examples might inspire rule-minded individuals to lower their moral standards and increase dishonesty. The comparison of ex-ante and ex-post effects of disclosing the information about the moral behavior of others offers an intriguing venue for future research.

4.6 Appendix

4.6.1 Recalled Ethical and Unethical Acts

Typical examples of ethical behavior include helping friends or relatives with moving into a new apartment or transporting: "I have helped my friend to move in. He had a couple of heavy pieces of furniture and many boxes...", "I have offered my car to a friend for moving out, although it could have been damaged", "I have spent my whole week-end to help a friend to move in, although I had to

write my home assignment", "I have transported a box with bottles of water with my car for a friend, since she lives on a hill and has to invest 20 min walking..", "A friend called me at 3.00 in the morning and told me he missed the last train..". Taking additional workload (in a shared flat or in the studies) is another common example of what participants consider ethical: "I have cooked a dinner for my roommate who called me after a long working day and cleaned the kitchen afterwards, although that should be her duty", "I have cleaned the apartment of my friend and cooked for him, although as his birthday party was. He could better concentrate on the studies...", Several participants mentioned their volunteering experience: "I volunteer as a tutor for children of different classes and help them to have good grades...", "I volunteer and help families with severely diseased children...Children from all parts of Germany have the opportunity to come to Jena and stay in a house provided by our organization next to the children clinic..", "..I have helped to load the container with medical equipment for a hospital in Malawi..." Only a few participants mention direct monetary donations: "Just before the experiment I gave a change to a street-musician for his music. I find it important to support these people as long as their music makes the atmosphere in the city more cheerful, and everybody can benefit from it."

The largest class of "unethical" examples are rejections of a help request with faked reasons: "I was driving to my appointment (app. 13 km). To pick up a friend I would have to make just 2km extra circle. However, I didn't and told her it was too far away for me", "I have told my parent in the morning that I could not sit with the baby since I am sick. Since I have not seen a friend for a long time, I wanted to chat with him instead...", "I wanted to go home on Thursday but I had to work until Friday. On Thursday evening I notified a boss I am sick, although it was not the case". Small direct

and indirect stealing is also a common response: "I have stolen a protein chocolate in a fitness studio. I had to wait long at the reception and got bored. I wanted to taste it. Besides, they charge far too much in that studio.", "In a supermarket I have replied to a question, whether I had 11 buns in my bag with "yes", although there were 12 of them. No risk, no fun", "We have decided to go to a birthday party and took a taxi to get there. The fare was a bit more expensive that we expected. We were six persons and had to pay 55 EUR. I have given only 6 EUR instead of 9 EUR and pretended and had no extra money with me.", "I have kept the money borrowed from a friend to pay my meal. She forgot that she gave me money and I did not remind her about it", "I have suggested a colleague not to come to the workplace as if there was not much to do (we have an hour pay scheme). There was a lot to do and I worked the whole day and earned more." Selfish decisions in previous experiments trigger negative feelings and are sometimes also classified as instances of "unethical" behavior by the participants.

4.6.2 Experimental Instructions

(Common text. Specific text for the baseline and treatment is marked separately).

Welcome and thank for participating in this experiment. Please switch off your mobile phones and remain silent. If you have any questions, please raise your hand and experimenter will answer you privately.

For your participation you will receive 3.00 EUR show-up fee. During the experiment you can earn additional money. Your ad-

ditional earnings depend on your own decisions, decisions of other participants as well as on chance. During the whole experiment your anonymity is guaranteed. Your additional earnings will be expressed in ECU (Experimental Currency Units) that will be converted into Euro at the end of the experiment at the following exchange rate: 10,00 ECU = 1,00€.

This experiment consists of three parts. The following instructions describe Part 1 and Part 3 of the experiment. The instructions for Part 2 will be presented to you on the screen once Part 1 is over.

At the beginning of the experiment all participants will be divided into two groups. You stay within the same group throughout the experiment. *3rd party treatment:* One participant in each group will be randomly selected to make a decision at the beginning of Part 2 of the experiment. This decision concerns all participants in his group but has no effect on himself. All the participants will be informed about this decision before Part 2 begins. In contrast to all other participants in his group, this randomly selected participant will not be asked to make any further decisions in the experiment.

Part 1: At the beginning of the experiment participants in both groups will be confronted with specific examples of everyday behavior. All participants will then be asked to provide one further example of such behavior, to recall their own recent activity and write down what they recall.

The information participants provide in this part of the experiment does not affect their earnings and will be kept strictly anony-

mous.

Part 2: Part 2 will be presented to you on the screen, once Part 1 is over. Part 2 is the only pay-off relevant part of the experiment.

Part 3: In the last part of the experiment you will be asked to solve certain number of 10 by 10 tables. These tables contain digits 0 and 1 in a randomized order. The example of such table is presented below:

0	0	1	1	0	1	1	0	0	1
1	1	1	0	1	1	1	1	1	0
1	1	1	0	1	1	0	0	1	1
0	0	1	0	0	0	1	0	0	1
1	1	0	1	0	1	1	1	1	0
0	0	1	0	1	0	1	0	1	1
0	1	1	0	1	0	0	0	0	0
0	0	1	1	1	0	1	0	1	1
1	1	0	0	0	0	0	1	0	0
0	0	1	1	1	1	1	0	1	1

Once you see a randomly generated table you have to count how often digit 1 appears in the table and enter this number (with ± 1 tolerance). If the number your enter is wrong, you will receive a new table until you give the right answer. The table is then counted as solved.

Baseline: The number of tables to solve is fixed for each group: one group will be asked to solve **1 table** and the other group will be asked to solve **8 tables.** You will be informed whether you (and the other participants) of your group will have to solve 1 or 8 tables

after Part 1 of the experiment. Actual solving of tables takes place only in Part 3 of the experiment.

3rd party treatment: After Part 1 of the experiment one participant in each group will be randomly selected. This participant decides whether other members of his group will have to solve **1 table** or **8 tables**. This participant himself does not have to solve any tables and makes no decisions in Part 2 of the experiment.

Once you have correctly solved the defined number of tables, you will have the opportunity to surf in the internet until all the other participants are ready with the task.

Before the payment begins, all participants will be asked to answer several questions about the experiment.

This will conclude the experiment.

If you have read the instructions carefully and do not have any further questions, click on 'Continue' to start Part 1 of the experiment.

Part 2: Instructions (presented on the screen after Part 1)

In this part of the experiment you will be matched with another participant from you group. In each pair there are two roles: Participant A and Participant B. The roles will be assigned randomly.

Participant A receives an endowment of 100 ECU and decides how much from this amount he would like to give to Participant B. The amount specified by Participant A will be deducted from his account and transferred to Participant B.

Before the participants know if they receive the role of Participant A or Participant B, both of them will be asked to make a decision as Participant A.

The computer will then assign the role of Participant A to one of the participants and implement his decision. The participants will be informed about their roles, the decision of (selected) Participant A and their earnings.

If you have read the instructions for Part 2 carefully and do not have any further questions, please click on 'Continue'.

Instructions for "Judges" (presented on the screen after Part 1)

You have an endowment of **100 ECU** and can decide, how many tables the participants in your group will have to solve. The default that we set is that the participants in your group at the end of the experiment will have to solve **1 table (8 tables)**.

You can change the number of tables they have to solve to 8 (1) table(s).

If you change the default number of tables to solve, you loose your endowment of 100 ECU.

If you DO NOT change the default the number of tables to solve, you keep your endowment of 100 ECU.

Do you want to change the default number of tables?

Chapter 5

Discussion

5.1 Conclusion and methodological remarks

People's preferences for pro-social and ethical behavior are influenced by a number of situational factors present in everyday social interactions. The experiments in this thesis identify some of these environmental features affecting the resulting degree of people's ethicality and pro-sociality.

The findings of these experiments support the importance of 'the power of the situation' in the domain of pro-social behavior and call for the necessary pre-caution of categorizing people into types (for example, by measuring their social value orientation Murphy *et al.* (2011)) to understand the outcomes of social interactions. In other words, the interplay between the environment and the person produces no 'sinners' or 'saints' but rather people's 'successes' and 'failures' to sacrifice their interest to the benefit of others.

Interpretation of the specific findings of the experiments, however, requires certain precautions.

First important question is the degree of external validity. How

does participant's willingness to delegate lies in a laboratory in Jena, as shown in one of the projects in this thesis, translate into the behavior of an employee in a particular bank in Hamburg or of an expert in an anti-corruption governmental agency in Ukraine?

Of course, Ukrainian anti-corruption agency is very different from the experimental laboratory in Jena in terms of size, actors, or incentives. The question is how systematic are those differences. If one expects that a certain feature of the anti-corruption institution would make actors behave differently in comparison to students in Jena, nothing prevents a researcher from isolating this effect in the lab. Any ex-ante claims about low external validity of the laboratory experiment remain an unsupported hypothesis. Thus, a director of anti-corruption office in Kyiv, can treat the willingness to delegate in our three-person game by students in Jena as a reasonable working expectation of the behavior of his larger team of deputies, experts, and social activists.

Another important precaution one should take is that the empirical analysis in this thesis is based on observation rather than on direct measurement of people's motivation. The work done in this thesis can be viewed as an attempt to link psychological studies that help to identify the environmental factors affecting pro-social and ethical behavior and economic experiments that explore people's behavior under real monetary incentives. Although we control for a number of factors and impose monetary stimuli, we can only say that certain manipulation produces a particular pattern of behavior and cannot identify internal processes leading to this behavioral change. We observe, for example, a larger incidence of lying when delegation is possible but we are not able to tell much about the changes in decision making processes or internal motivations caused by that delegation possibility. In this respect, a user of this

experimental work would benefit from complementing the findings in this thesis by the evidence from relevant studies in psychology and neuroeconomics.

The last specific methodological issue concerns the qualitative discrepancy in the findings of the 2nd and the 3rd project of the thesis. In the study on delegation of lies, we have observed compensatory behavior: lying was associated with larger donations in the subsequent task. In the study on third-party influence on one's moral behavior, we have found the evidence for consistent behavior: an unethical act was followed by lower donations. The divergence of these findings is likely to be driven by two methodological differences between the studies. First, remembering an unethical deed (chapter 4) and actually performing it (chapter 3) are two different matters. Although the priming used in chapter 4 is shown to evoke compensatory behavior in other experiments, it might still function as a mere exposition known to enhance consistency (see Mazar & Zhong (2010) for a discussion). Second, one-shot (chapter 4) and repeated (chapter 3) interaction might produce different reactions from the participants. Repeated unethical action might accumulate more negative feelings and call for a larger compensation. The fact that in the delegation experiment compensatory tendencies appear only in later rounds supports this intuition.

5.2 Implications and venues for future research

In the first project the possibility to remain ignorant to ex-post information about recipient's endowment decreased giving in a dictator game with incomplete information. Notably, ex-post information disclosure is only relevant when participants had the chance to choose if they want to reveal recipient's endowment. Exogenous

manipulation of information disclosure does not seem to affect dictator's giving. These findings add to other experimental evidence on increased selfishness under ex-ante ignorance (Dana *et al.*, 2007; Lazear *et al.*, 2012).

Although the possibility to remain ignorant is viewed negatively because it increases selfishness and inequality, it might have positive effects in settings where selfish behavior coincides with socially optimal behavior. One can expect, for example, that less information about the intentions of others might decrease the incidence of negative reciprocity known to dramatically damage efficiency in many social interactions. Gueth *et al.* (2014) allow responders in ultimatum game to costly acquire the information about the true pie size. They observe that only a few responders invest into information acquisition, and moreover the rate of destructions by responders for unknown pie sizes are lower. If negative reciprocity is partially driven by negative social emotions, avoiding certain information might be strategic for people with strong emotional regulation, as emotional self-regulation theory suggests (Koole, 2009). If that is true, (dis-) incentivizing people for acquiring such information might improve the results of more complex social interactions such as multi-stage bargaining or contest-like competitions.

The experiment in chapter 2 is the zero-sum situation, i.e., a transfer from a dictator could by no means change the total social welfare. It might be interesting, however, to investigate dictator's reactions towards uncertainty and recipient's requests for help in cases when transfers could potentially increase efficiency. The results of the experiment also provide important insights into the behavior of recipients who are rarely completely passive in the real world in contrast to a stylized dictator game. In particular, the recipients seem to underestimate the influence of their request on

dictator's giving and thus fail to follow an 'optimal' begging strategy. The closer analysis of recipients' readiness to ask for help might be relevant for improving the design of welfare programs.

In the second project, we have observed that delegation opportunity might increase incidence of lying. Participants who tell the truth when no delegation is possible, delegate more often that those who are ready to lie themselves. These findings signal that hierarchical structures with delegation possibility can be more vulnerable to dishonesty. In positive domain, in contrast, as other experiments mentioned in the chapter show, people might underdelegate benevolent activities. Therefore, a policy maker should consider impeding the opportunity for delegation of socially undesirable or unethical activities and stimulate delegation where employment of an agent increases social welfare.

Since both social and self-evaluation of the behavior were present in the experiment (receivers learned if delegation took place) it is important to explore the relative strength of these two factors. If the social evaluation enhances people's willingness to delegate socially blamed activities like lying, would one observe more delegation if the audience becomes larger? Another important aspect of the study is the link between one's responsibility behind the transgression and the willingness to compensate for it with a subsequent action. We have found that delegating the decision to lie is associated with more generous transfers in the subsequent dictator game. It suggests that lower objective responsibility behind unethical actions does not necessarily imply lower willingness to compensate. Future studies can benefit from direct investigation of how the degree of responsibility behind one's action enters into licensing/balancing behavior. In the project we have focused on the unethical behavior, however, one should try to extent the analysis for the

ethical domain. Will the fact the one shares the responsibility for a good deed with someone else reduce licensing behavior?

In the third project we address the issue of how one's own ethical behavior interact with the moral behavior of others. We understand that people rarely behave in insolation but often encounter examples that guide their behavior in a given environment: a firm, a school, or a family. We argue that identifying people's mindset (outcome or rule-based) is helpful in understanding how people react on an example of behavior they see. We suggest that this distinction is useful in a variety of applications. An employer that builds up a team to work together on a project or a tax authority disclosing identity of tax evaders in public should be interested in how the information about others affects the behavior of the actors they target. We found that rule-minded individuals react asymmetrically strong to negative action of others: unethically primed individuals significantly reduced their transfers in a subsequent dictator game. In turn, outcome-based participants were symmetrically insensitive to observing behavior of others. These findings suggest that negative examples might provoke an increase in unethical behavior among rule-minded individuals, whereas positive examples do not suffice to boost morality.

All three experiments in this thesis are organized along the idea of self- evaluation of one's behavior. Apart from an observation by anonymous another participant the social context was deliberately excluded. Judgment by others, however, is an important feature of any social order. Therefore, to understand ethical and pro-social behavior in a richer context, it is important to investigate how others evaluate and react to (mis-)behavior of actors.

Ndodjang *et al.* (2013) show that responders were willing to accept lower offers from proposers who agreed to volunteer for a

public project. These findings suggest that higher moral account earns license from others affected by the decision of the actors.

Social order, however, requires that not only affected second parties but also third parties unaffected by the current interaction intervene (see Fehr & Fischbacher (2004) for a discussion). In this respect, the evaluation and costly reaction towards the (un-)ethical behavior of others by unaffected third-parties in a dynamic context offers another intriguing venue for future research.

The knowledge of how people adjust their ethical judgment and reaction to the history of actor's choices opens another possibility to enhance socially optimal behavior.

Bibliography

Abbink, K. & Herrmann, B. (2011). The Moral Costs Of Nastiness. *Economic Inquiry*, *49*(2), pp. 631–633.

Aimone, J.A. & Houser, D. (2013). Harnessing the benefits of betrayal aversion. *Journal of Economic Behavior & Organization*, *89*, pp. 1–8.

Akerlof, G.A. & Kranton, R.E. (2000). Economics and Identity. *The Quarterly Journal of Economics, MIT Press*, *115*(3), pp. 715–753.

Akerlof, G. & Kranton, R. (2010). Identity Economics. *The Economists' Voice*, *7*(2), pp. 1–3.

Alm, J., Bloomquist, K.M. & McKee, M. (2013). *When You Know Your Neighbor Pays Taxes: Information, Peer Effects, and Tax Compliance*, Working Papers 13-22, Department of Economics, Appalachian State University.

Andreoni, J. (1990). Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving? *Economic Journal*, *100*(401), pp. 464–77.

Andreoni, J. & Bernheim, B.D. (2009). Social Image and the 50-50 Norm: A Theoretical and Experimental Analysis of Audience Effects. *Econometrica*, *77*(5), pp. 1607–1636.

Andreoni, J. & Rao, J.M. (2011). The power of asking: How com-

munication affects selfishness, empathy, and altruism. *Journal of Public Economics, Elsevier*, *95*(7-8), pp. 513–520.

Baron, J. & Ritov, I. (2004). Omission bias, individual differences, and normality. *Organizational Behavior and Human Decision Processes*, *94*(2), pp. 74–85.

Bartling, B. & Fischbacher, U. (2012). Shifting the blame: On delegation and responsibility. *Review of Economic Studies*, *79*(1), pp. 67–87.

Battigalli, P., Charness, G. & Dufwenberg, M. (2013). Deception: The role of guilt. *Journal of Economic Behavior & Organization*, *93*, pp. 227–232.

Battigalli, P. & Dufwenberg, M. (2007). Guilt in games. *American Economic Review*, *97*(2), pp. 170–176.

Benabou, R. & Tirole, J. (2011). Identity, morals and taboos: Beliefs as assets. *Quarterly Journal of Economics*, *126*(2), pp. 805–855.

Bohnet, I. & Zeckhauser, R. (2004). Trust, risk and betrayal. *Journal of Economic Behavior & Organization*, *55*(4), pp. 467–484.

Buchanan, J. (1975). *The Samaritan's dilemma. In: Altruism, Morality and Economic Theory*, Russell Sage Foundation, New York, pp. 71– 85.

Casal, S., Güth, W., Jia, M. & Ploner, M. (2012). Would you mind if I get more? An experimental study of the envy game. *Journal of Economic Behavior & Organization*, *84*(3), pp. 857–865.

Casal, S. & Mittone, L. (2014). *Social Esteem versus Social Stigma: the role of anonymity in an income reporting game*, CEEL Working Papers 1401, Cognitive and Experimental Economics Laboratory, Department of Economics, University of Trento, Italia.

Charness, G. & Dufwenberg, M. (2006). Promises and partnership. *Econometrica*, *74*(6), pp. 1579–1601.

Charness, G. & Rabin, M. (2002). Understanding Social Preferences With Simple Tests. *The Quarterly Journal of Economics*, *117*(3), pp. 817–869.

Clot, S., Grolleau, G. & Ibanez, L. (2013). Self-Licensing and Financial Rewards: Is Morality For Sale? *Economics Bulletin*, *33*(3), pp. 2298–2306.

Clot, S., Grolleau, G. & Ibanez, L. (2014). Smug alert! exploring self-licensing behavior in a cheating game. *Economic Letters*, *123*, pp. 191–194.

Coffman, L.C. (2011). Intermediation reduces punishment (and reward). *American Economic Journal: Microeconomics*, *3*(4), pp. 77–106.

Cojoc, D. & Stoian, A. (2014). Dishonesty and charitable behavior. *Experimental Economics*, *17*, pp. 717–732.

Conway, P. & Peetz, J. (2012). When does feeling moral actually make you a better person? conceptual abstraction moderates whether past moral deeds motivate consistency or compensatory behavior. *Personality and Social Psychology Bulletin*, *38*(7), pp. 907–919.

Cornellissen, G., Bashshur, M.R., Rode, J. & Menestrel, M.L. (2013). Rules or consequences? the role of ethical mind-sets in moral dynamics. *Psychologicla Science*, *24*(4), pp. 241–262.

Dana, J., Weber, R. & Kuang, J. (2007). Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness. *Economic Theory*, *33*(1), pp. 67–80.

DeScioli, P., Christner, J. & Kurzbahn, R. (2011). The omission strategy. *Psychological Science*, *22*, pp. 442–446.

Drugov, M., Hamman, J. & Serra, D. (2014). Intermediaries in corruption: An experiment. *Experimental Economics*, pp. 78–99.

Eisenkopf, G. & Fischbacher, U. (2011). *Doing well by doing good*

- *or doing better by delegating?*, Twi research paper series, Thurgauer Wirtschaftsinstitut, Universität Konstanz.

Engel, C. (2011). Dictator games: a meta study. *Experimental Economics*, *14*(4), pp. 583–610.

Engel, C., Kurschilgen, M.J. & Kube, S. (2011). *Can we manage first impressions in cooperation problems?*, Working paper 2011-05, Max Planck Institute for Research on Collective Goods.

Erat, S. (2013). Avoiding lying: The case of delegated deception. *Journal of Economic Behavior and Organization*, *93*, pp. 273–278.

Erat, S. & Gneezy, U. (2012). White Lies. *Management Science*, *58*(4), pp. 723–733.

Fehr, E. & Fischbacher, U. (2004). Third party punishment and social norms. *Evolution and Human Behavior*, *25*, pp. 63–87.

Fehr, E. & Schmidt, K.M. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, *114*(3), pp. 817–868.

Festinger, L. (1962). Cognitive dissonace. *Scientific American*, *207*, pp. 93–107.

Fischbach, A., Dahr, R. & Zhang, Y. (2006). Subgoals as substitutes or complements: the role of goal accessebility. *Journal of Personality and Social Psychology*, *91*, pp. 232–242.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, *10*(2), pp. 171–178.

Fischbacher, U. & Föllmi-Heusi, F. (2013). Lies In Disguise—An Experimental Study On Cheating. *Journal of the European Economic Association*, *11*(3), pp. 525–547.

Freedman, J. & Fraser, S. (1966). Compliance without pressure: the foot-in-the door technique. *Journal of Personality and Social Psychology*, *4*, pp. 195–202.

Gino, F., Norton, M.I. & Ariely, D. (2010). The counterfeit self: The deceptive costs of faking it. *Psychological Science*, *21*(5), pp. 712–720.

Gneezy, A., Imas, A., Brown, A., Nelson, L.D. & Norton, M.D. (2012a). Paying to be nice: Consistency and costly prosocial behavior. *Management Science*, *58*(1), pp. 179–187.

Gneezy, U. (2005). Deception: The role of consequences. *American Economic Review*, *95*(1), pp. 384–394.

Gneezy, U., Imas, A. & Madarasz, K. (2012b). Conscience accounting: Emotional dynamics and social behavior. *Theoretical economics paper series , STICERD, London, UK.*

Greiner, B. (2004). An online recruitment system for economic experiments, in (K. Kremer & V. Macho, eds.), *Forschung und wissenschaftliches Rechnen*, vol. 63, of *GWDG Bericht*, pp. 79–93, Göttingen: Ges. für Wiss. Datenverarbeitung.

Gueth, W., Huck, S. & Ockenfels, P. (1996). Two-Level Ultimatum Bargaining with Incomplete Information: An Experimental Study. *Economic Journal*, *106*(436), pp. 593–604.

Gueth, W., Levati, V., Nardi, C. & Sorappera, I. (2014). An ultimatum game with multidimentional response strategies. *Jena Economic Research Papers*, (018).

Hamman, J.R., Loewenstein, G. & Weber, R.A. (2010). Self-interest through delegation: An additional rationale for the principal-agent relationship. *American Economic Review*, *100*(4), pp. 1826–46.

Harbring, C. & Irlenbusch, B. (2009). *Sabotage in Tournaments: Evidence from a Laboratory Experiment*, IZA Discussion Papers 4205, Institute for the Study of Labor (IZA).

Hayashi, A.T. (2013). Ocassionally libertarian: Experimental evi-

dence of self-serving omission bias. *Journal of Law, Economics and Organization*, *29*, pp. 711–733.

Houser, D., Vetter, S. & Winter, J. (2012). Fairness and cheating. *European Economic Review*, *56*(8), pp. 1645–1655.

Jordan, J., Mullen, E. & Murnighan, J.K. (2011). Striving for the moral self: The effects of recalling past moral actions on future moral behavior. *Personality and Social Psychology Bulletin*, *37*(5), pp. 707–713.

Kahn, U. & Dhar, R. (2006). Licensing effect in consumer choice. *Journal of Marketing Research*, *43*, pp. 259–266.

Kartik, N. (2009). Strategic communication with lying costs. *Review of Economic Studies*, *26*, pp. 1359–1395.

Khalmetski, K., Ockenfels, A. & Werner, P. (2013). *Surprising Gifts - Theory and Laboratory Evidence*, Working Paper Series in Economics, University of Cologne, Department of Economics 61, University of Cologne, Department of Economics.

Konow, J. (2000). Fair Shares: Accountability and Cognitive Dissonance in Allocation Decisions. *American Economic Review*, *90*(4), pp. 1072–1091.

Koole, S. (2009). The psychology of emotional regulation: An intergative review. *Cognition and Emotion*, *23*, pp. 4–41.

Langenbach, P. (2014). The values of ex-ant and ex-post communication in dictator games. *Max Planck Institute for Research on Collective Goods, Working paper 2014/7*.

Lazear, E.P., Malmendier, U. & Weber, R.A. (2012). Sorting in Experiments with Application to Social Preferences. *American Economic Journal: Applied Economics*, *4*(1), pp. 136–63.

Loewenstein, G. (2000). Emotions in economic theory and economic behavior. *American Economic Review*, *90*(2), pp. 426–432.

Lovelady, S. (2014). *Experimental Regret Aversion*, Ph.D. thesis, University of Warwick.

Matthey, A. & Regner, T. (2011). Do I Really Want to Know? A Cognitive Dissonance-Based Explanation of Other-Regarding Behavior. *Games*, *2*(1), pp. 114–135.

Mazar, N., Amir, O. & Ariely, D. (2008). The dishonesty of honest people: A theory of self- concept maintance. *Journal of Marketing Research*, *45*(6), pp. 633–644.

Mazar, N. & Zhong, C.B. (2010). Do green products make us better people? *Psychological Science*, *21*(4), pp. 494–498.

Meritt, A.T., Efron, D. & Monin, B. (2010). Moral self-licensing: When being good frees us to be bad. *Social and Personality Psychology Compass*, *4*, pp. 344–357.

Monin, B. & Miller, D.T. (2001). Moral credentials and the expressions of prejudice. *Journal of Personality and Social Psychology*, *81*, pp. 344–357.

Murphy, R.O., Ackerman, K.A. & Handgraaf, M.J.J. (2011). Measuring social value orientation. *Judgment and Decision Making*, *6*(8), pp. 771–781.

Ndodjang, P., Grolleau, G. & Ibanez, L. (2013). Do previous good deeds to a third party make people more tolerant of bad deeds against them? An experimental investigation. *Economics Letters*, *121*(3), pp. 364–368.

Nisan, M. & Horenczyk, G. (1990). Moral balance: the effect of prior behavior on the decicion in moral context. *British Journal of Social Psychology*, *29*, pp. 29–42.

Ockenfels, A. & Bolton, G.E. (2000). ERC: A Theory of Equity, Reciprocity, and Competition. *American Economic Review, American Economic Association*, *90*(1), pp. 166–193.

Patt, A. & Zeckhauser, R. (2000). Action bias and environmental decisions. *Journal of Risk and Uncertainty*, *21*, pp. 45–72.

Rauhut, H. (2013). Beliefs about lying and spreading of dishonesty: Undetected lies and their constructive and destructive social dynamics in dice experiments. *PLoS One*, *8*, pp. 1–8.

Royzman, E.B. & Baron, J. (2002). The peferences for indirect harm. *Social Justice Research*, *15*(2), pp. 165–185.

Sachdeva, S., Iliev, R. & Menin, D.S. (2009). Sinning saints and saintly sinners: The paradox of moral self-regulation. *Psychological Science*, *20*, pp. 523–528.

Schwarzwald, J., Bizman, A. & Raz, M. (1983). The foot-in-the-door paradigm: Effects of second request size on donation probability and donor generosity. *Personallity and Social Psychology Bulletin*, *9*(3), pp. 443–450.

Shalvi, S., Dana, J., Handgraaf, M.J. & De Dreu, C.K. (2011). Justified ethicality: Observing desired counterfactuals modifies ethical perceptions and behavior. *Organizational Behavior and Human Decision Processes*, *115*(2), pp. 181–190.

Spranca, M., Minsk, E. & Baron, J. (1991). Omission and comission in judgement and choice. *Journal of Experimental Social Psychology*, *27*, pp. 76–105.

Sutter, M. (2009). Deception through telling the truth?! experimental evidence from individuals and teams. *Economic Journal*, *119*(534), pp. 47–60.

Thöni, C. (2014). Inequality aversion and antisocial punishment. *Theory and Decision*, *76*(4), pp. 529–545.

Trope, Y. & Liberman, N. (2003). Temporal construal. *Psychological Review*, *110*(3), pp. 403–421.

Vanberg, C. (2008). Why do people keep their promises? an exper-

imental test of two explanations -super-1. *Econometrica*, *76*(6), pp. 1467–1480.

Wilson, J. & Kelling, G. (1982). Broken windows: The police and neighborhood safety. *Atlantic Monthly*, *127*, pp. 29–38.

Winschel, E. & Zahn, P. (2014). *When Ignorance is Bliss - Information Asymmetries Enhance Prosocial Behavior in Dictator Games*, CESifo Working Paper Series 4750, CESifo Group Munich.

Yamamori, T., Kato, K., Kawagoe, T. & Matsui, A. (2008). Voice matters in a dictator game. *Experimental Economics*, *11*(4), pp. 336–343.

Cirriculum Vitae

SERHIY KANDUL
Graduate School IMPRS 'Uncertainty'
Kahlaische str. 10, 07745
Jena, Germany
kandul@econ.mpg.de
office 251.2
+49 3641 686 673

1. DATE OF BIRTH: November 06, 1984

2. CITIZENSHIP: Ukrainian

3. CIVIL STATUS: Married

4. EDUCATION:

- 2006-2008 MA in Economics, Economics Education and Research Consortium, Kyiv- Mohyla Academy. Kyiv, Ukraine.
- 2001-2006 BA in International Economics, Dnipropetrovsk National University, Ukraine.
- 1991-2001 Secondary education certificate, Tukali and Semenivka Schools, Semenivskiy district, Poltavska oblast, Ukraine.

5. PRESENT EMPLOYER AND POSITION:

- Doctoral Student at University of Jena Germany
  (since November 2011)

6. SCHOLARSHIP/FELLOWSHIPS:

- August 2008 Teaching Fellowship, CERGE-EI and Czech National Academy of Sciences, Prague, Czech Republic
- November 2011-December 2014, MPI of Economics, Doctoral Fellowship

7. MISCELLANEOUS:

Languages: English (good), German (B2), Russian (very good), Ukrainian (mother tongue)
Computer skills: Fundamentals of R, z-Tree, LaTeX
Hobbies: Knopfakkordeon B-Griff (bajan).

Erklärung nach 4 Abs. 1 PromO:

Hiermit erkläre ich:

1. dass mir die geltende Promotionsordnung bekannt ist;
2. dass ich die Dissertation selbst angefertigt, keine Textabschnitte eines Dritten oder eigener Prüfungsarbeiten ohne Kennzeichnung ubernommen und alle von mir benutzten Hilfsmittel, persönlichen Mitteilungen und Quellen in meiner Arbeit angegeben habe;
3. dass ich bei der Auswahl und Auswertung des Materials sowie bei der Herstellung des Manuskriptes keine unzulässige Hilfe in Anspruch genommen habe;
4. dass ich nicht die Hilfe eines Promotionsberaters in Anspruch genommen habe und dass Dritte weder unmittelbar noch mittelbar geldwerte Leistungen von mir fur Arbeiten erhalten haben, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen;
5. dass ich die Dissertation noch nicht als Prüfungsarbeit fur eine staatliche oder andere wissenschaftliche Prüfung eingereicht habe;
6. dass ich nicht die gleiche, eine in wesentlichen Teilen ahnliche oder eine andere Abhandlung bei einer anderen Hochschule bzw. anderen Fakultat als Dissertation eingereicht habe

Jena, den Serhiy Kandul